

Master 2 : EDP dispersives

Sylvie Benzoni¹

23 février 2010

¹Université Claude Bernard Lyon 1, benzoni@math.univ-lyon1.fr

L'objectif de ce cours est de présenter un panorama de propriétés analytiques et algébriques de quelques équations aux dérivées partielles (EDP) modélisant des phénomènes de propagation d'ondes dispersives. Les domaines d'applications des EDP étudiées sont variés : vibrations, vagues, transitions de phase, mécanique quantique, etc.

Chapitre I

EDP dispersives linéaires

1 Exemples et définitions de base

On s'intéresse dans ce premier chapitre à des EDP d'évolution linéaires, dans lesquelles l'inconnue est une fonction (voire une distribution)

$$\begin{aligned} \mathbf{u} : \mathbb{R}^+ \times \mathbb{R}^d &\rightarrow \mathbb{C}^N \\ (t, \mathbf{x}) &\mapsto \mathbf{u}(t, \mathbf{x}). \end{aligned}$$

(Tout au long de ces notes, les caractères gras seront utilisés pour des valeurs vectorielles ou matricielles.) Dans toutes les applications considérées, t représentera une variable temporelle et les composantes de \mathbf{x} dans \mathbb{R}^d des variables spatiales. Sauf mention contraire, les EDP linéaires étudiées seront à coefficients constants. On considèrera principalement des EDP d'ordre 1 en temps, c'est-à-dire de la forme

$$\partial_t \mathbf{u} + \sum_{\alpha} \mathbf{A}_{\alpha} \partial_{\alpha} \mathbf{u} = 0,$$

où les α sont des multi-indices $\alpha = (\alpha_1, \dots, \alpha_d)$ de « longueur » $|\alpha| = \sum_j \alpha_j$ (supposée majorée par un entier m), les \mathbf{A}_{α} sont des matrices (réelles ou complexes) $N \times N$, et

$$\partial^{\alpha} := \partial_1^{\alpha_1} \dots \partial_d^{\alpha_d}$$

est un opérateur de dérivation d'ordre $|\alpha|$. Si $N = 1$ on parle d'EDP scalaire. En voici quelques exemples.

Équation de transport : $\partial_t u + \mathbf{a} \cdot \nabla u = 0$, $\mathbf{a} \in \mathbb{R}^d$.

Équation d'Airy : $\partial_t u + a \partial_x u + k \partial_{xxx}^3 u = 0$, $a, k \in \mathbb{R}$.

Équation de Schrödinger : $i \partial_t u + h \Delta u = 0$, $h \in \mathbb{R}^{+*}$,

cette dernière pouvant également être vue comme un système d'inconnue u à valeurs dans \mathbb{R}^2 . On étudiera également des équations d'ordre 2 en temps, construites autour de l'équation des ondes.

Équation des ondes : $\partial_{tt}^2 u - c^2 \Delta u = 0$, $c \in \mathbb{R}^+$.

Équation de Klein-Gordon : $\partial_{tt}^2 u - c^2 \Delta u + k u = 0$, $c \in \mathbb{R}^+$, $k \in \mathbb{R}^*$.

Équation de Boussinesq : $\partial_{tt}^2 u - c^2 \partial_{xx}^2 u + k \partial_{xxxx}^4 u = 0$, $c \in \mathbb{R}^+$, $k \in \mathbb{R}^*$,

Les paramètres présents dans ces équations proviennent de constantes physiques. Toutes ces équations ont comme point commun de modéliser des phénomènes de propagation d'ondes, et l'on verra que la plupart sont *dispersives*, en un sens que l'on précisera. Une vague « définition » est la suivante : *une EDP est dispersive si elle propage des fréquences différentes à des vitesses différentes*. La suite du chapitre permettra notamment de donner un sens à cette formulation.

Définition I.1

Une onde plane est une fonction $u : (t, \mathbf{x}) \rightarrow u(t, \mathbf{x})$ de la forme

$$u(t, \mathbf{x}) = U(\boldsymbol{\xi} \cdot \mathbf{x} + \tau t),$$

où $\boldsymbol{\xi} \in (\mathbb{R}^d)'$ et $\tau \in \mathbb{R}$.

Par définition, si $\mathbf{v} \in \mathbb{R}^3$ appartient à l'hyperplan $\{\mathbf{v}; \boldsymbol{\xi} \cdot \mathbf{v} = -\tau\}$, et si u est une onde plane comme ci-dessus, on a pour tout $(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^d$,

$$u(t, \mathbf{x}) = u(0, \mathbf{x} - \mathbf{v}t).$$

Autrement dit, la valeur de u à l'instant t sur un hyperplan $\{\mathbf{x}; \boldsymbol{\xi} \cdot \mathbf{x} = \alpha\}$ (qui ne dépend pas du point sur cet hyperplan) est entièrement déterminée par sa valeur à l'instant $t = 0$ sur l'hyperplan $\{\mathbf{x}; \boldsymbol{\xi} \cdot \mathbf{x} = \alpha + \tau t\}$.

Dans ce qui suit on considère (momentanément) t comme une variable ordinaire, également notée x_0 , et les multi-indices comportent a priori une composante α_0 pour les dérivées par rapport à t .

Définition I.2

On appelle symbole d'une EDP linéaire à coefficients constants

$$\sum_{|\alpha| \leq m} \mathbf{A}_\alpha \partial_\alpha \mathbf{u} = 0,$$

la fonction polynômiale

$$\begin{aligned} p : (\mathbb{R}^{1+d})' &\rightarrow \mathbb{C}^{N \times N} \\ \boldsymbol{\xi} &\mapsto p(\boldsymbol{\xi}) := \sum_{|\alpha| \leq m} (i\boldsymbol{\xi})^\alpha \mathbf{A}_\alpha, \quad \text{où } \boldsymbol{\xi}^\alpha := \xi_0^{\alpha_0} \cdots \xi_d^{\alpha_d}. \end{aligned}$$

Le symbole principal est la fonction polynômiale

$$p_m : \boldsymbol{\xi} \mapsto p(\boldsymbol{\xi}) := \sum_{|\alpha|=m} (i\boldsymbol{\xi})^\alpha \mathbf{A}_\alpha,$$

en supposant que les matrices \mathbf{A}_α ci-dessus ne sont pas toutes nulles (auquel cas il faudrait diminuer m). Lorsque $p = p_m$, on dit que l'EDP est homogène. Un vecteur $\boldsymbol{\xi} \in (\mathbb{R}^{1+d})'$ est dit caractéristique pour l'EDP si

$$\det p_m(\boldsymbol{\xi}) = 0.$$

Parmi les exemples ci-dessus les seules EDP homogènes sont l'équation de transport, pour laquelle $p_m(\boldsymbol{\xi}) = i(\xi_0 + \sum_{j=1}^d a_j \xi_j)$, et l'équation des ondes, pour laquelle $p_m(\boldsymbol{\xi}) =$

$c^2\|\xi\|^2 - \xi_0^2$. Ces deux équations appartiennent à la classe des EDP hyperboliques, dont on rappelle ci-dessous la définition.

Définition I.3

Une EDP linéaire à coefficients constants de symbole principal p_m est dite hyperbolique si

- le vecteur $(1, 0, \dots, 0)$ n'est pas caractéristique,
- pour tout $\xi \in (\mathbb{R}^{1+d})'$, la fonction polynômiale $z \mapsto \det p_m(\xi_0 + z, \xi_1, \dots, \xi_d)$ a toutes ses racines réelles.

Plus généralement, pour $\eta \in (\mathbb{R}^{1+d})'$ l'EDP est dite hyperbolique dans les directions transverses à l'hyperplan $\{\eta \cdot (t, \mathbf{x}) = 0\}$ de \mathbb{R}^{1+d} si

- le vecteur η n'est pas caractéristique,
- pour tout $\xi \in (\mathbb{R}^{1+d})'$, la fonction polynômiale $z \mapsto \det p_m(\xi + z\eta)$ a toutes ses racines réelles.

Par définition, une EDP hyperbolique admet au moins une direction caractéristique, $\zeta = (\tau, \xi_1, \dots, \xi_d) \in (\mathbb{R}^{1+d})' \setminus \{0\}$. À cette direction correspondent des ondes planes solutions de l'EDP si elle est homogène : il suffit en effet de choisir un vecteur \mathbf{V} dans le noyau de $p_m(\zeta)$ et une fonction scalaire $f : \theta \mapsto f(\theta)$ arbitraire ; l'onde plane

$$u : (t, \mathbf{x}) \mapsto f(\xi \cdot \mathbf{x} + \tau t) \mathbf{V}$$

est solution de l'EDP. (Ici ξ désigne à nouveau un élément de $(\mathbb{R}^d)'$, ses composantes étant naturellement (ξ_1, \dots, ξ_d) .)

Pour les équations non homogènes il n'est pas aussi facile de trouver des solutions sous forme d'ondes planes. On peut cependant rechercher des ondes planes particulières, périodiques en $\xi \cdot \mathbf{x} + \tau t$, que l'on appelle alors *phase*, et plus spécialement sous la forme d'ondes planes progressives monochromatiques (OPPM) :

$$u : (t, \mathbf{x}) \mapsto u(t, \mathbf{x}) = e^{i(\xi \cdot \mathbf{x} + \tau t)}.$$

On obtient alors la *relation de dispersion* de l'EDP, qui s'exprime à l'aide du symbole (complet) sous la forme

$$\boxed{\det p(\tau, \xi_1, \dots, \xi_d) = 0.}$$

Dans une OPPM comme ci-dessus, la période temporelle est $T = 2\pi/|\tau|$ et $\omega = -\tau$ (c'est une convention) est la *pulsation*.

2 Dispersion

Afin d'amener progressivement les notions fondamentales liées aux phénomènes dispersifs, on commence par se placer en une dimension d'espace.

2.1 Dimension un

Dans ce qui suit, on considère une EDP linéaire scalaire, de symbole p , en dimension 1 d'espace. Si $(\xi, \omega) \in \mathbb{R}^2$ est tel que $p(-\omega, \xi) = 0$, l'OPPM

$$u : (t, x) \mapsto u(t, x) = e^{i(\xi x - \omega t)}$$

est solution de l'EDP. Le réel ξ est appelé *nombre d'onde* (il faudrait plutôt dire nombre d'onde angulaire) : si λ désigne la *longueur d'onde*, c'est-à-dire par définition la période spatiale, on a $\xi = 2\pi/\lambda$.

Comme on l'a déjà vu, si le symbole p est de plus homogène, *toutes* les ondes planes

$$u : (t, x) \mapsto u(t, x) = U(\xi x - \omega t),$$

où $p(-\omega, \xi) = 0$ et U est une fonction *arbitraire*, sont des solutions de l'EDP. Si $\xi \neq 0$ (ce qui est automatique pour une EPD hyperbolique par exemple, puisque $(1, 0)$ n'est pas caractéristique), on voit que ces ondes planes se propagent à la vitesse $v = \omega/\xi$, une quantité homogène de degré 0 et donc indépendante de ξ : par exemple pour l'équation de transport, $v = a$, et pour l'équation des ondes $v = \pm c$.

Qu'en est-il pour les équations non homogènes ? Une OPPM de nombre d'onde $\xi \neq 0$ et de pulsation $\omega = W(\xi)$ se propage à une vitesse $W(\xi)/\xi$ qui dépend en général de ξ : c'est le sens de l'énoncé « elle propage des fréquences différentes à des vitesses différentes » dans la vague définition que nous avons donnée plus haut. Nous allons maintenant pouvoir préciser cette définition, en introduisant les notions de *vitesse de phase* et de *vitesse de groupe*.

Afin de motiver les définitions de ces notions, nous allons étudier le comportement de solutions plus générales que les OPPM, en l'occurrence des « paquets d'ondes », obtenus en « superposant » des OPPM. En effet, si on a une famille de solutions de l'EDP sous forme d'OPPM de pulsation $\omega = W(\xi)$, la fonction

$$u : (t, x) \mapsto u(t, x) = \frac{1}{2\pi} \int \widehat{u}_0(\xi) e^{i(\xi x - W(\xi)t)} d\xi,$$

où u_0 est (par exemple) dans la classe de Schwartz, est solution de l'EDP et vérifie la condition initiale $u(0, x) = u_0(x)$. Ci-dessus $\widehat{u}_0(\xi)$ désigne la transformée de Fourier de u_0 au point ξ : on vérifie ainsi que $u(0, x) = u_0(x)$ grâce à la *formule d'inversion de Fourier* (pour laquelle on pourra se reporter à l'appendice). Le fait que u soit effectivement solution de l'EDP provient de la « dérivation sous le signe somme », parfaitement autorisée si u_0 et toutes ses dérivées sont intégrables, ce qui est le cas pour $u_0 \in \mathcal{S}(\mathbb{R})$.

La *méthode de la phase stationnaire* (voir par exemple [4, p.136]) permet d'étudier le comportement de $u(t, x)$ à x/t fixé lorsque $t \rightarrow +\infty$. Supposons qu'il existe un unique ζ tel que $W'(\zeta) = x/t$, et que $W''(\zeta) \neq 0$ (c'est le cas par exemple si W est strictement convexe). Alors

$$u(t, x) \sim \frac{\widehat{u}_0(\zeta) e^{-i \operatorname{sgn} W''(\zeta) \pi/4}}{\sqrt{4\pi t |W''(\zeta)|}} e^{i(\zeta x - W(\zeta)t)} =: \varphi(t, x) =: A(t, x) e^{i\theta(t, x)}.$$

Lorsque l'on fait varier (t, x) , le nombre d'onde « local » $\zeta = \zeta(t, x)$ défini implicitement par

$$W'(\zeta(t, x)) = x/t,$$

est une solution autosimilaire de l'équation hyperbolique

$$\partial_t \zeta + W'(\zeta) \partial_x \zeta = 0.$$

La vitesse $W'(\zeta)$ est ce que l'on appellera la vitesse de groupe de l'EDP, ou plus précisément de la famille des OPPM de pulsation donnée par $\omega = W(\xi)$. Par ailleurs, on voit que la phase $\theta(t, x) := \zeta(t, x)x - W(\zeta(t, x))t$ vérifie

$$\partial_x \theta = \zeta, \quad \partial_t \theta = -W(\zeta),$$

et elle se propage par conséquent à la vitesse $W(\zeta)/\zeta$, au sens où θ est constant le long des trajectoires de l'EDO de

$$\frac{dX}{dt} = \frac{W(\zeta(t, X))}{\zeta(t, X)}.$$

Définition I.4

Soit une famille d'OPPM dont la pulsation est donnée en fonction du nombre d'onde par $\omega = W(\xi)$. On appelle vitesse de groupe de cette famille la fonction

$$v_g : \xi \in \mathbb{R}^* \mapsto W'(\xi),$$

et vitesse de phase la fonction

$$v_\varphi : \xi \in \mathbb{R}^* \mapsto W(\xi)/\xi.$$

Si une EDP linéaire monodimensionnelle admet comme solutions une famille d'OPPM avec une vitesse de phase et une vitesse de groupe différentes, l'EDP est dite dispersive. On appelle lignes de groupe les courbes caractéristiques issues de 0 de

$$\partial_t \zeta + W'(\zeta) \partial_x \zeta = 0,$$

(qui sont en fait des rayons $x/t = \text{constante}$), et lignes de phase les trajectoires de $X' = v_\varphi(\zeta(t, X))$, où ζ est donnée implicitement par $W'(\zeta(t, x)) = x/t$.

On constate immédiatement que vitesse de groupe et vitesse de phase coïncident lorsque W est homogène de degré 1, ce qui est le cas si l'EDP est homogène. (On a alors $W(\xi) = |\xi|W(\text{sgn}\xi)$.) Ainsi l'équation de transport et l'équation des ondes (mono-dimensionnelle) ne sont *pas dispersives*, ce qui est cohérent avec la remarque faite plus haut : pour ces EDP, toutes les ondes planes se propagent à la même vitesse (au signe près pour l'équation des ondes). En revanche les autres EDP données en exemple sont dispersives.

Remarque I.1

La définition ci-dessus peut apparaître restrictive, dans la mesure où certaines équations présentent une combinaison d'effets dispersifs et d'effets *diffusifs*. C'est le cas par exemple de l'équation obtenue en ajoutant un terme d'ordre 2 à l'équation d'Airy

$$\partial_t u - \mu \partial_{xx} u + k \partial_{xxx}^3 u = 0, \quad (\mu > 0).$$

Cette EDP n'admet pas d'OPPM (elle admet en fait des « ondes » *diffusives*, de la forme $u(t, x) = e^{i(\xi x + \tau t)}$ avec $\text{Im}\tau > 0$, et donc exponentiellement décroissantes quand $t \rightarrow +\infty$), mais selon les valeurs respectives des paramètres μ et k , elle peut être vue comme une perturbation de l'équation d'Airy ou une perturbation de l'équation de la chaleur

$$\partial_t u = \mu \partial_{xx} u,$$

qui est le modèle de base des phénomènes de *diffusion* (ici en une dimension).

On a vu que la vitesse de phase était la vitesse de propagation de la phase θ pour

$$\varphi(t, x) = A(t, x) e^{i\theta(t, x)}, \quad \text{avec } \theta(t, x) = \zeta(t, x)x - W(\zeta(t, x))t \quad \text{et } W'(\zeta(t, x)) = x/t,$$

$$A(t, x) := \frac{\widehat{u}_0(\zeta(t, x)) e^{-i \operatorname{sgn} W''(\zeta(t, x)) \pi/4}}{\sqrt{2\pi t |W''(\zeta(t, x))|}}.$$

Quant à la vitesse de groupe, on a vu que c'était la vitesse de propagation du *nombre d'onde local* ζ . Elle s'interprète aussi comme la vitesse de propagation de l'amplitude A , ou plus exactement de $|A|^2$. En effet, supposons $W''(\xi) > 0$ quel que soit ξ . Alors pour $x_1 < x_2$,

$$\begin{aligned} \int_{x_1}^{x_2} |A(t, x)|^2 dx &= \frac{1}{2\pi} \int_{x_1}^{x_2} \frac{|\widehat{u}_0(\zeta(t, x))|^2}{t W''(\zeta(t, x))} dx = \\ &= \frac{1}{2\pi} \int_{\xi_1}^{\xi_2} |\widehat{u}_0(\xi)|^2 d\xi \end{aligned}$$

par changement de variables tel que $x = W'(\xi)t$. Si l'on fixe ξ_1 et ξ_2 , les points $x_1 = W'(\xi_1)t$ et $x_2 = W'(\xi_2)t$ sont les positions à l'instant t des lignes de groupe associées aux nombre d'ondes locaux ξ_1 et ξ_2 , et l'égalité ci-dessus montre que $\int_{x_1(t)}^{x_2(t)} |A(t, x)|^2 dx$ est constante au cours du temps. La décroissance de A en $1/\sqrt{t}$ est liée au fait que les lignes de groupe se séparent en $\mathcal{O}(t)$.

L'exposé précédent, en terme de paquets d'ondes, est fondé sur la transformation de Fourier, et par conséquent spécifique aux équations linéaires à coefficients constants. On peut donner une autre approche de la notion de vitesse de groupe, qui ne passe pas par la transformation de Fourier, et s'avère applicable à des équations plus générales (à coefficients variables, voire non linéaires). Elle est fondée sur ce que l'on appelle l'« *optique géométrique* » (domaine mathématique inspiré de la physique des rayons lumineux). Le point de départ de cette approche est la recherche de solutions dépendant d'un petit paramètre ε , exprimées en fonction des variables lentes $\tilde{x} = \varepsilon x$ et $\tilde{t} = \varepsilon t$. Il s'agit plus précisément de chercher des solutions sous forme de trains d'onde, dont l'amplitude est en $\mathcal{O}(1)$ et la phase en $\mathcal{O}(1/\varepsilon)$,

$$u(t, x) = A(\varepsilon t, \varepsilon x; \varepsilon) e^{iS(\varepsilon t, \varepsilon x)/\varepsilon}.$$

À titre d'exemple, considérons l'équation d'Airy,

$$\partial_t u + a \partial_x u + k \partial_{xxx}^3 u = 0,$$

qui devient dans les variables lentes

$$\partial_{\tilde{t}} u + a \partial_{\tilde{x}} u + \varepsilon^2 k \partial_{\tilde{x}\tilde{x}\tilde{x}}^3 u = 0.$$

Dans la suite du calcul on omet les tildas par commodité. On cherche donc une solution de

$$\partial_t u + a \partial_x u + \varepsilon^2 k \partial_{xxx}^3 u = 0,$$

sous la forme

$$u(t, x) = A(t, x; \varepsilon) e^{iS(t, x)/\varepsilon}.$$

En identifiant les puissances de ε (en supposant l'amplitude A analytique par rapport à ε), on obtient formellement l'équation

$$(I.1) \quad \partial_t S + a \partial_x S - k (\partial_x S)^3 = 0,$$

appelée *équation eikonale*. C'est une équation de *Hamilton-Jacobi* : elle s'écrit en effet

$$H(t, x, \partial_t S, \partial_x S) = 0$$

pour le Hamiltonien

$$H(t, x, \tau, \xi) = \tau + a\xi - k\xi^3$$

(qui ne dépend en fait pas de (t, x) ici). Une équation de Hamilton-Jacobi en général se « résout » par la *méthode des (bi)caractéristiques*, reposant sur le système d'équations différentielles ordinaires (EDO)

$$(I.2) \quad \begin{cases} x' = \frac{\partial H}{\partial \xi}, \\ \xi' = -\frac{\partial H}{\partial x}, \\ t' = \frac{\partial H}{\partial \tau}, \\ \tau' = -\frac{\partial H}{\partial t}, \end{cases}$$

et l'observation que H est constant le long des courbes intégrales de ce système. On peut ainsi calculer S solution de $H(t, x, \partial_t S, \partial_x S) = 0$ en intégrant le système (I.2) à partir d'un point $(t_0, x_0, \tau_0, \xi_0)$ tel que $H(t_0, x_0, \tau_0, \xi_0) = 0$, puis en intégrant l'équation $S' = \tau t' + \xi x'$. Dans l'exemple choisi, le système (I.2) se réduit à

$$\begin{cases} x' = a - 3k\xi^2, \\ \xi' = 0, \\ t' = 1, \\ \tau' = 0, \end{cases}$$

et ses courbes intégrales sont des droites paramétrées par t . On obtient ainsi une solution $S = S(t, x; \xi)$ de (I.1) pour chaque valeur de $\xi = \xi_0$. Par construction, elle est telle que $\partial_x S = \xi$ et on voit en dérivant (I.1) que

$$\partial_t(\partial_\xi S) + (a - 3k\xi^2)\partial_x(\partial_\xi S) = 0.$$

Autrement dit, ξ est le nombre d'onde local associé à la phase S , et les points où cette phase est stationnaire, c'est-à-dire où $\partial_\xi S = 0$, se propagent à la vitesse de groupe (ici $a - 3k\xi^2$). Ceci est un fait général. En effet, pour une EDP linéaire d'ordre 1 (pour simplifier) à coefficients réels

$$\partial_t u + \sum_{\ell=0}^m a_\ell \partial_x^\ell u = 0,$$

qui devient après changement d'échelle,

$$\varepsilon \partial_t u + \sum_{\ell=0}^m \varepsilon^\ell a_\ell \partial_x^\ell u = 0,$$

l'équation eikonale est

$$i\partial_t S + \sum_{\ell=0}^m i^\ell a_\ell (\partial_x S)^\ell = 0,$$

c'est-à-dire $p(\partial_t S, \partial_x S) = 0$, où p est le symbole de l'EDP. Ainsi l'équation eikonale est une équation de Hamilton-Jacobi pour le Hamiltonien $H = p/i$! En particulier, la relation de dispersion n'est autre que $H(-\omega, \xi) = 0$. Elle est à coefficients réels si l'EDP ne contient que des dérivées d'ordre impair, ce que l'on supposera. En posant $W(\xi) = \sum_{\ell} i^{\ell-1} a_{\ell} \xi^{\ell}$, la relation de dispersion s'écrit encore $\omega = W(\xi)$. On a $H(\tau, \xi) = \tau + W(\xi)$, et l'équation eikonale est

$$\partial_t S + W(\partial_x S) = 0.$$

Le système d'EDO associé (I.2) s'écrit

$$\begin{cases} x' = W'(\xi), \\ \xi' = 0, \\ t' = 1, \\ \tau' = 0, \end{cases}$$

et la projection sur le plan $\{(t, x)\}$ de ses courbes intégrales fournit précisément ce que l'on a appelé les lignes de groupe. Pour chaque valeur de ξ on trouve ainsi une solution $S = S(t, x; \xi)$ de l'équation eikonale telle que $\partial_x S = \xi$, et par dérivation de cette équation, on a

$$\partial_t(\partial_{\xi} S) + W'(\xi) \partial_x(\partial_{\xi} S) = 0.$$

Le raisonnement ci-dessus a l'avantage de s'étendre à des EDP à coefficients variables, ce qui donne un Hamiltonien dépendant vraiment de (t, x) , et fournit encore une interprétation de la vitesse de groupe.

2.2 Dimension quelconque

Définition I.5

La variété caractéristique est l'ensemble

$$\mathcal{C} = \{(\tau, \xi) \in (\mathbb{R} \times \mathbb{R}^d)'; \det p(\tau, \xi) = 0\}.$$

Au voisinage d'un point $(\tau, \xi) \in \mathcal{C}$ où elle est régulière, il existe une fonction W telle que \mathcal{C} ait pour équation $\tau + W(\xi) = 0$. La co-normale à \mathcal{C} en un tel point s'identifie au vecteur de \mathbb{R}^{1+d} dont la première composante est 1 et les autres sont les dérivées partielles de W par rapport à ξ . Ces dernières définissent sans ambiguïté la vitesse de groupe $\mathbf{v}_g : (W(\xi), \xi) \in \mathcal{C} \mapsto \nabla_{\xi} W$ pour l'EDP. Par ailleurs, une vitesse de phase est une fonction $\mathbf{v}_{\varphi} : (W(\xi), \xi) \in \mathcal{C} \mapsto \mathbf{v}_{\varphi}$ telle que

$$\xi \cdot \mathbf{v}_{\varphi}(\xi) = W(\xi).$$

On dira que l'EDP est dispersive s'il existe des points réguliers de sa variété caractéristique où la vitesse de groupe n'est pas une vitesse de phase.

On remarque que si W est homogène de degré 1, ce qui est le cas si l'EDP est homogène, la vitesse de groupe est une vitesse de phase. C'est le cas en particulier pour l'équation des ondes.

Pour une EDP multi-dimensionnelle, l'optique géométrique fournit encore une équation de Hamilton-Jacobi comme équation eikonale. Si la relation de dispersion s'écrit

$$\omega = W(\boldsymbol{\xi}),$$

l'équation eikonale est

$$\partial_t S + W(\nabla_{\mathbf{x}} S) = 0.$$

Elle se résout en considérant le système d'EDO

$$\begin{cases} \mathbf{x}' = \nabla_{\boldsymbol{\xi}} W(\boldsymbol{\xi}), \\ \boldsymbol{\xi}' = 0, \\ t' = 1, \\ \tau' = 0 \end{cases}$$

et en intégrant $S' = \tau t' + \boldsymbol{\xi} \cdot \mathbf{x}'$. Sans surprise, on obtient pour chaque valeur de $\boldsymbol{\xi}$, l'expression « naturelle » de la phase

$$S(t, \mathbf{x}; \boldsymbol{\xi}) = \boldsymbol{\xi} \cdot \mathbf{x} - W(\boldsymbol{\xi})t$$

(à une constante près).

nom EDP	relation dispersion	vitesse de groupe	vitesse de phase
transport	$\omega = \boldsymbol{\xi} \cdot \mathbf{a}$	$\mathbf{v}_g = \mathbf{a}$	$\boldsymbol{\xi} \cdot (\mathbf{v}_\varphi - \mathbf{a}) = 0$
Airy	$\omega = a\xi - k\xi^3$	$v_g = a - 3k\xi^2$	$v_\varphi = a - k\xi^2$
Schrödinger	$\omega = h\ \boldsymbol{\xi}\ ^2$	$\mathbf{v}_g = 2h\boldsymbol{\xi}$	$\boldsymbol{\xi} \cdot (\mathbf{v}_\varphi - h\boldsymbol{\xi}) = 0$
ondes	$\omega = \pm c\ \boldsymbol{\xi}\ $	$\mathbf{v}_g = \pm c \frac{\boldsymbol{\xi}}{\ \boldsymbol{\xi}\ }$	$\boldsymbol{\xi} \cdot \left(\mathbf{v}_\varphi \mp c \frac{\boldsymbol{\xi}}{\ \boldsymbol{\xi}\ } \right) = 0$
Klein-Gordon	$\omega = \pm \sqrt{k + c^2\ \boldsymbol{\xi}\ ^2}$	$\mathbf{v}_g = \pm c^2 \frac{\boldsymbol{\xi}}{\sqrt{k + c^2\ \boldsymbol{\xi}\ ^2}}$	$\boldsymbol{\xi} \cdot \mathbf{v}_\varphi = \pm \sqrt{k + c^2\ \boldsymbol{\xi}\ ^2}$
Boussinesq	$\omega = \pm \sqrt{c^2\xi^2 + k\xi^4}$	$v_g = \pm \frac{c^2\xi + 2k\xi^3}{\sqrt{c^2\xi^2 + k\xi^4}}$	$v_\varphi = \pm \sqrt{c^2 + k\xi^2}$

TAB. I.1 – Exemples de relations de dispersion, vitesses de groupe et vitesses de phase.

3 Formules de résolution

Avant d'aborder des théorèmes plus ou moins abstraits concernant le problème de Cauchy, nous allons tâcher d'obtenir des formules de résolution « explicites », ce qui est possible pour les EDP linéaires à coefficients constants posées dans tout l'espace grâce notamment à la transformation de Fourier. Le but est d'obtenir des majorations fines du *noyau de Green* de ces EDP : pour une EDP linéaire d'ordre un en temps, le noyau de Green, ou encore la *solution fondamentale*, est par définition une application $t \mapsto G_t \in \mathcal{S}'(\mathbb{R}^d)$, solution (faible) de l'EDP pour $t > 0$ et telle que $\lim_{t \rightarrow 0} G_t = \delta$ (la masse de

Dirac en $\mathbf{x} = 0$) ; lorsque G_t s'avère être une fonction de \mathbf{x} pour tout $t > 0$, on parle de *fonction de Green*. Lorsqu'on connaît le noyau de Green, on peut résoudre le problème de Cauchy pour toute donnée initiale u_0 qui se convole avec G_t quel que soit $t > 0$: pour un tel u_0 , la fonction $t \mapsto u(t) := u_0 *_{\mathbf{x}} G_t$ est solution de l'EDP et vérifie $\lim_{t \rightarrow 0} u(t) = u_0$. La recherche du noyau de Green se fait précisément en résolvant le problème de Cauchy pour une donnée initiale arbitraire u_0 .

Équation d'Airy Si u est solution de l'EDP

$$\partial_t u + a \partial_x u + k \partial_{xxx}^3 u = 0,$$

la fonction $\widehat{u} = \mathcal{F}u$ (où la transformation de Fourier \mathcal{F} agit sur la variable spatiale seulement) est solution de l'EDO

$$\partial_t \widehat{u} + i(a\xi - k\xi^3) \widehat{u} = 0,$$

et par conséquent

$$\widehat{u}(t, \xi) = \widehat{u}_0(\xi) e^{(-ia\xi + ik\xi^3)t},$$

où $u_0 := u(0, \cdot)$. Donc, si u_0 est suffisamment localisée en espace (par exemple à support compact),

$$u(t, \cdot) = u_0 * \mathcal{F}^{-1}(\xi \mapsto e^{(-ia\xi + ik\xi^3)t}).$$

La fonction analytique $\xi \mapsto e^{(-ia\xi + ik\xi^3)t}$ étant oscillante (mais tout de même bornée!), sa transformée de Fourier inverse est *a priori* une distribution tempérée. La multiplication par $e^{-ia\xi t}$ en Fourier correspondant simplement à une translation par $-at$ en variables spatiale, il suffit de calculer la transformée de Fourier inverse de $\xi \mapsto e^{ik\xi^3 t}$. Il n'y en fait pas de formule explicite, mais dans le cas $kt = 1/3$ cette transformée inverse connue sous le nom de fonction d'Airy

$$\text{Ai}(x) := \frac{1}{2\pi} \int e^{i\xi x + \frac{1}{3}\xi^3} d\xi = \frac{1}{\pi} \int_0^{+\infty} \cos(\xi x + \frac{1}{3}\xi^3) d\xi.$$

Par la *méthode du col* [4, pp. 279–289] (voir aussi [15, p. 442]), on montre que

$$\text{Ai}(x) \sim \frac{1}{2\sqrt{\pi}} x^{-1/4} e^{-\frac{2}{3}x^{3/2}}, \quad x \rightarrow +\infty,$$

$$\text{Ai}(x) \sim \frac{1}{\sqrt{\pi}} |x|^{-1/4} \sin(\frac{2}{3}|x|^{3/2} + \pi/4), \quad x \rightarrow -\infty.$$

Par suite, un petit changement de variables montre que $\mathcal{F}^{-1}(\xi \mapsto e^{ik\xi^3 t})$ est la fonction $x \mapsto \frac{1}{\sqrt[3]{3kt}} \text{Ai}(x/\sqrt[3]{3kt})$. Donc finalement, $u(t, \cdot) = u_0 * K_t$, le noyau de Green K_t étant donné par

$$K_t(x) = \frac{1}{\sqrt[3]{3kt}} \text{Ai}\left(\frac{x - at}{\sqrt[3]{3kt}}\right).$$

La fonction Ai étant bornée, on voit en particulier que

$$\|K_t\|_{L^\infty(\mathbb{R})} = \mathcal{O}(t^{-1/3})$$

lorsque $t \rightarrow +\infty$. Et par ailleurs, à $t > 0$ fixé, on a $K_t \in L^4(\mathbb{R})$ d'après les estimations ci-dessus de Ai.

Équation de Schrödinger Si u est solution de

$$i\partial_t u + h\Delta u = 0$$

et vaut u_0 à $t = 0$, par transformation de Fourier on trouve

$$\widehat{u}(t, \boldsymbol{\xi}) = \widehat{u}_0(\boldsymbol{\xi}) e^{ih\|\boldsymbol{\xi}\|^2 t},$$

d'où $u(t, \mathbf{x}) = (u_0 * H_t)(\mathbf{x})$ avec

$$H_t(\mathbf{x}) := \left(\frac{1}{4h\pi t} \right)^{d/2} e^{-i d \pi / 4} e^{i \frac{\|\mathbf{x}\|^2}{4h t}}, \quad t > 0.$$

La justification du calcul de $\mathcal{F}^{-1}(\boldsymbol{\xi} \mapsto e^{ih\|\boldsymbol{\xi}\|^2 t})$ (au sens des distributions) est fondée sur un argument de fonction de variable complexe et la formule pour le *noyau de la chaleur*, donnée pour $\alpha > 0$ par

$$G_t(\mathbf{x}) := \mathcal{F}^{-1}(\boldsymbol{\xi} \mapsto e^{-\alpha\|\boldsymbol{\xi}\|^2 t}) = \left(\frac{1}{4\alpha\pi t} \right)^{d/2} e^{-\frac{\|\mathbf{x}\|^2}{4\alpha t}}, \quad t > 0.$$

On voit que pour $u_0 \in L^1(\mathbb{R})$, la fonction $u(t) : \mathbf{x} \mapsto (u_0 * H_t)(\mathbf{x})$ est continue pour tout $t > 0$. Si u_0 est carrément à support compact, $u(t)$ est de classe \mathcal{C}^∞ pour tout $t > 0$. Ceci est l'*effet régularisant local* de l'équation de Schrödinger. (En comparaison, l'équation de la chaleur est globalement régularisante : pour tout $u_0 \in L^\infty(\mathbb{R}^d)$, l'application $\mathbf{x} \mapsto (u_0 * G_t)(\mathbf{x})$ est de classe \mathcal{C}^∞ pour tout $t > 0$.) En contrepartie de cet effet régularisant local, $u(t)$ cesse instantément d'être à support compact : ceci témoigne d'un phénomène de propagation à vitesse infinie dans l'équation de Schrödinger, attribué à la *dispersion* (alors que la propagation à vitesse infinie dans l'équation de la chaleur est attribuée à la *diffusion*).

Par comparaison avec l'équation d'Airy, on a une meilleure décroissance en temps, puisque pour $d = 1$,

$$\|H_t\|_{L^\infty(\mathbb{R})} = \mathcal{O}(t^{-1/2})$$

lorsque $t \rightarrow +\infty$. Ceci est lié au fait que « les basses fréquences sont plus rapides » (voir le tableau I.1) : les vitesses de propagation sont de l'ordre de ξ dans l'équation de Schrödinger, et ξ^2 dans l'équation d'Airy (en supposant $a = 0$, ce qui revient à faire un changement de référentiel). En revanche, la régularisation en espace est moins bonne que pour l'équation d'Airy, puisque le noyau H_t n'a pas de propriété de décroissance à l'infini (quand $x \rightarrow \pm\infty$) : ceci est lié au fait que les hautes fréquences (responsables des défauts de régularité) sont moins rapides.

Équation des ondes Pour les EDP d'ordre deux en temps, une « donnée initiale » consiste en deux fonctions (notées ci-après u_0 et u_1) déterminant non seulement les valeurs de la solution à $t = 0$ mais aussi les valeurs de sa dérivée partielle par rapport à t . Dans le cas de l'équation des ondes

$$\partial_{tt}^2 u - c^2 \Delta u = 0,$$

on obtient par transformation de Fourier en espace

$$\partial_{tt}^2 \widehat{u} + c^2 \|\boldsymbol{\xi}\|^2 \widehat{u} = 0.$$

Cette équation différentielle ordinaire d'ordre deux se résout en

$$\widehat{u}(t, \boldsymbol{\xi}) = \widehat{u}_0(\boldsymbol{\xi}) \cos(ct\|\boldsymbol{\xi}\|) + \widehat{u}_1(\boldsymbol{\xi}) \frac{\sin(ct\|\boldsymbol{\xi}\|)}{c\|\boldsymbol{\xi}\|},$$

et ainsi

$$u(t, \cdot) = u_0 * K_t^0 + u_1 * K_t^1,$$

$$K_t^0 = \mathcal{F}^{-1}(\boldsymbol{\xi} \mapsto \cos(ct\|\boldsymbol{\xi}\|)), \quad K_t^1 = \mathcal{F}^{-1}\left(\boldsymbol{\xi} \mapsto \frac{\sin(ct\|\boldsymbol{\xi}\|)}{c\|\boldsymbol{\xi}\|}\right).$$

Attention cependant, les noyaux K_t^0 et K_t^1 sont ici des distributions (à symétrie radiale) : ceci est lié à la vitesse finie de propagation (isotrope) dans l'équation des ondes, notion que l'on va préciser avec le principe de Huyghens.

Dimension $d = 1$. On vérifie aisément que

$$K_t^0 = \frac{1}{2}(\delta_{ct} + \delta_{-ct}),$$

en exprimant pour $f \in \mathcal{S}$,

$$\langle \widehat{K}_t^0, f \rangle = \frac{1}{2}(\widehat{f}(ct) + \widehat{f}(-ct)) = \int \cos(ctx) f(x) dx,$$

et

$$K_t^1 = \frac{1}{2c} \mathbf{1}_{[-c|t|, c|t|]},$$

où $\mathbf{1}_{[-c|t|, c|t|]}$ désigne la fonction caractéristique de l'intervalle $[-c|t|, c|t|]$, car alors pour $\xi \neq 0$,

$$\widehat{K}_t^1(\xi) = \frac{1}{2c} \int_{-c|t|}^{c|t|} e^{-i\xi x} dx = \frac{\sin(ct\xi)}{c\xi}.$$

On obtient ainsi une formule explicite pour la solution de l'équation des ondes, la *formule de d'Alembert* :

$$u(t, x) = \frac{1}{2}(u_0(x + ct) + u_0(x - ct)) + \frac{1}{2c} \int_{x-c|t|}^{x+c|t|} u_1(y) dy.$$

(Il y a diverses autres manières de l'obtenir, sans passer par la transformation de Fourier !)

Dimension $d = 3$. On peut montrer (par la méthode dite des moyennes sphériques par exemple), la *formule de Kirchhoff* :

$$u(t, \mathbf{x}) = \partial_t \left(\frac{1}{4\pi c^2 t} \int_{\|\mathbf{y}-\mathbf{x}\|=ct} u_0(\mathbf{y}) d\mathbf{y} \right) + \frac{1}{4\pi c^2 t} \int_{\|\mathbf{y}-\mathbf{x}\|=ct} u_1(\mathbf{y}) d\mathbf{y}.$$

Il existe une formule analogue en toute dimension impaire (voir par exemple [14, p. 72]).

On voit sur cette formule que si les données initiales sont concentrées dans la boule rayon R de centre \mathbf{x} , la solution au point (x, t) est nulle pour tout $t > R/c$. Ceci est connu comme le *principe de Huyghens*.

Dimension $d = 2$ On peut montrer (en utilisant la formule de Kirchhoff pour des fonctions qui ne dépendent pas de la troisième variable !) que

$$u(t, \mathbf{x}) = \partial_t \left(\frac{1}{2\pi c} \int_{\|\mathbf{y}-\mathbf{x}\|^2 \leq c^2 t^2} \frac{u_0(\mathbf{y})}{\sqrt{c^2 t^2 - \|\mathbf{y}-\mathbf{x}\|^2}} d\mathbf{y} \right)$$

$$+ \frac{1}{2\pi c} \int_{\|\mathbf{y}-\mathbf{x}\|^2 \leq c^2 t^2} \frac{u_1(\mathbf{y})}{\sqrt{c^2 t^2 - \|\mathbf{y}-\mathbf{x}\|^2}} d\mathbf{y}.$$

On remarque cette fois que même avec des données initiales concentrées dans une boule autour d'un point, la solution en ce point ne « s'éteint » jamais.

Équation de Klein-Gordon On voit naturellement apparaître une distinction entre le « bon cas » ($k > 0$) et le « mauvais cas » ($k < 0$), pour lequel l'équation différentielle du second ordre

$$\partial_{tt}^2 \hat{u} + (k + c^2 \|\boldsymbol{\xi}\|^2) \hat{u} = 0$$

admet des solutions exponentiellement croissantes. Si l'on se place dans le cas $k > 0$, on obtient

$$\hat{u}(t, \boldsymbol{\xi}) = \hat{u}_0(\boldsymbol{\xi}) \cos t \sqrt{k + c^2 \|\boldsymbol{\xi}\|^2} + \hat{u}_1(\boldsymbol{\xi}) \frac{\sin t \sqrt{k + c^2 \|\boldsymbol{\xi}\|^2}}{\sqrt{k + c^2 \|\boldsymbol{\xi}\|^2}},$$

ce qui ne donne pas grand chose par transformation de Fourier inverse ! Une astuce consiste à remarquer que la fonction

$$v : (t, x_1, \dots, x_d, y) \in \mathbb{R} \times \mathbb{R}^{d+1} \mapsto v(t, x_1, \dots, x_d, y) := e^{i\sqrt{k}y/c} u(t, x_1, \dots, x_d)$$

est solution de l'équation des ondes en dimension $d + 1$,

$$\partial_{tt}^2 v - c^2 \Delta v = 0,$$

si u est solution de l'équation de Klein-Gordon en dimension d .

Équation de Boussinesq Là encore il y a le « bon cas » ($k > 0$) et le « mauvais cas » ($k < 0$). Nous en reparlerons au prochain paragraphe. Dans le bon, la formule de résolution s'écrit en variable de Fourier

$$\hat{u}(t, \boldsymbol{\xi}) = \hat{u}_0(\boldsymbol{\xi}) \cos t \sqrt{k \|\boldsymbol{\xi}\|^4 + c^2 \|\boldsymbol{\xi}\|^2} + \hat{u}_1(\boldsymbol{\xi}) \frac{\sin t \sqrt{k \|\boldsymbol{\xi}\|^4 + c^2 \|\boldsymbol{\xi}\|^2}}{\sqrt{k \|\boldsymbol{\xi}\|^4 + c^2 \|\boldsymbol{\xi}\|^2}}.$$

4 Estimations *a priori*

4.1 Conservation de l'énergie

Toutes les EDP citées en exemple dans le paragraphe 1 ont la propriété de conserver une « énergie ». Pour les équations d'ordre 1, il s'agit simplement de la norme L^2 .

Proposition I.1

Soit $u \in \mathcal{C}^1(0, T; L^2(\mathbb{R}^d)) \cap \mathcal{C}(0, T; H^m(\mathbb{R}^d))$ solution de l'équation de transport ($m = 1$), de l'équation de Schrödinger ($m = 2$), ou de l'équation d'Airy ($m = 3, d = 1$). Alors pour tout $t \in [0, T]$,

$$\|u(t)\|_{L^2} = \|u(0)\|_{L^2}.$$

Démonstration : Elle repose sur le calcul de

$$\frac{d}{dt} \int |u(t, \mathbf{x})|^2 d\mathbf{x} = 2\operatorname{Re} \int \overline{u(t, \mathbf{x})} \partial_t u(t, \mathbf{x}) d\mathbf{x}$$

et d'intégrations par parties.

Pour l'équation de transport on a

$$\frac{d}{dt} \int |u(t, \mathbf{x})|^2 d\mathbf{x} = -2\operatorname{Re} \int \overline{u(t, \mathbf{x})} \mathbf{a} \cdot \nabla u(t, \mathbf{x}) d\mathbf{x} = \int |u(t, \mathbf{x})|^2 \operatorname{div} \mathbf{a} d\mathbf{x} = 0$$

si a est constant. En fait ce calcul donne aussi une estimation intéressante lorsque a est une fonction Lipschitzienne :

$$\|u(t)\|_{L^2} \leq \|u(0)\|_{L^2} e^{T\|\operatorname{div} \mathbf{a}\|_{L^\infty}}.$$

Ceci est ce que l'on appelle une *estimation a priori* (dans L^2).

Pour l'équation de Schrödinger on a

$$\frac{d}{dt} \int |u(t, \mathbf{x})|^2 d\mathbf{x} = 2\operatorname{Re} \int i h \overline{u(t, \mathbf{x})} \Delta u(t, \mathbf{x}) d\mathbf{x} = -2\operatorname{Re} \int i h |\nabla u(t, \mathbf{x})|^2 d\mathbf{x} = 0.$$

Enfin, pour l'équation d'Airy,

$$\begin{aligned} \frac{d}{dt} \int |u(t, x)|^2 dx &= -2\operatorname{Re} \int \overline{u(t, x)} (a \partial_x u(t, x) + k \partial_{xxx}^3 u(t, x)) dx = \\ &= \int |u(t, x)|^2 \partial_x a + k \partial_x ((\partial_x |u(t, x)|)^2) dx = 0 \end{aligned}$$

si a est constant (et k aussi : le cas d'un coefficient k variable est plus compliqué). ■

Pour les équations du second ordre il existe également des quantités conservées. Pour l'équation de Klein-Gordon (qui contient l'équation des ondes si on autorise $k = 0$), on constate que

$$\int ((\partial_t u)^2 + c^2 \|\nabla u\|^2 + k u^2)(t, \mathbf{x}) d\mathbf{x}$$

est conservée. Pour l'équation de Boussinesq, on trouve que

$$\int ((\partial_t u)^2 + c^2 (\partial_x u)^2 + k (\partial_{xx} u)^2)(t, x) dx$$

est conservée.

4.2 Estimations de dispersion

Nous allons dans ce paragraphe obtenir des estimations a priori supplémentaires, de type $L^{r'} \rightarrow L^r$, où selon la notation standard

$$\frac{1}{r} + \frac{1}{r'} = 1.$$

(On rappelle que le dual topologique de L^r est $L^{r'}$ pour tout $r \in [1, +\infty[$, et que le dual de L^∞ est strictement plus gros que L^1 .) Les estimations d'énergie du paragraphe

précédent traitent du cas $r = r' = 2$. Par ailleurs, la forme des noyaux des Green obtenus au paragraphe 3 montre les estimations suivantes dans le cas $r = +\infty$, $r' = 1$.

Pour l'équation d'Airy :

$$\|u(t)\|_{L^\infty(\mathbb{R})} \lesssim |t|^{-1/3} \|u_0\|_{L^1(\mathbb{R})}.$$

Pour l'équation de Schrödinger :

$$\|u(t)\|_{L^\infty(\mathbb{R}^d)} \lesssim |t|^{-d/2} \|u_0\|_{L^1(\mathbb{R}^d)}.$$

On peut ensuite procéder par *interpolation*, en utilisant le résultat classique suivant, qui est un cas particulier du théorème de Riesz-Thorin [2, p. 2!]).

Théorème I.1 (*Riesz-Thorin*)

Si T définit un opérateur linéaire continu de L^1 dans L^∞ , de norme M et aussi de L^2 dans L^2 , de norme N , alors il est continu de L^s dans L^r avec

$$\frac{1}{s} = 1 - \theta + \frac{\theta}{2}, \quad \frac{1}{r} = \frac{\theta}{2}, \quad 0 < \theta < 1,$$

et la norme de $T : L^s \rightarrow L^r$ est majorée par $M^{1-\theta} N^\theta$.

Corollaire I.1

Pour tout $r \in [2, +\infty[$, on a les estimations a priori suivantes, appelées estimations de dispersion pour l'équation d'Airy :

$$\|u(t)\|_{L^r(\mathbb{R})} \lesssim |t|^{-1/3+2/3r} \|u_0\|_{L^{r'}(\mathbb{R})},$$

et pour l'équation de Schrödinger (libre) :

$$\|u(t)\|_{L^r(\mathbb{R}^d)} \lesssim |t|^{-d/2+d/r} \|u_0\|_{L^{r'}(\mathbb{R}^d)}.$$

4.3 Estimations de de Strichartz

Depuis le travail fondateur de Strichartz (paru en 1977), on appelle inégalités de Strichartz des estimations a priori du type, pour les EDP d'ordre 1, $L^2 \rightarrow L^q(0, t; L^r(\mathbb{R}^d))$. Pour les EDP d'ordre 2 comme l'équation des ondes, on verra qu'il faut demander plus de régularité sur l'espace de départ (ce qui revient plus ou moins à considérer ces équations comme des *systèmes* d'ordre 1 en $\partial_t u$ et ∇u). Dans tous les cas, pour espérer obtenir des estimations (ou inégalités) de Strichartz, il ne faut pas que la variété caractéristique de l'EDP ait une composante *plate* : ceci exclut notamment l'équation de transport (dont la variété caractéristique est un hyperplan), et l'équation des ondes en dimension 1 (dont la « variété » caractéristique est composée de deux droites).

Les ingrédients pour obtenir ces estimations sont :

- des estimations (ou inégalités) de dispersion,
- un argument appelé T^*T ,
- l'inégalité de Hardy-Littlewood-Sobolev, aussi appelé théorème d'intégration fractionnaire.

Équations d'ordre 1

Des estimations de dispersion ont été obtenues au paragraphe 4.2 (Corollaire I.1) pour les équations d'Airy et de Schrödinger. Ces équations ont de plus en commun d'avoir des solutions données par un *groupe d'opérateurs unitaires*. Bien que cela ne soit pas en soi indispensable, nous allons utiliser ce groupe dans l'argument T^*T .

Rappelons tout d'abord les rudiments de la théorie des (semi-)groupes à un paramètre.

Semi-groupes fortement continus On appelle \mathcal{C}^0 -*semi-groupe* une famille d'opérateurs $(S_t)_{t \geq 0}$ telle que

- pour tout $t \geq 0$, S_t est un opérateur linéaire continu sur un espace de Banach \mathcal{X} ,
- pour tous $t, s \geq 0$, $S_t \circ S_s = S_{t+s}$,
- on a $S_0 = \text{id}$,
- et de plus, pour tout $\phi \in \mathcal{X}$,

$$\lim_{t \searrow 0} \|S_t \phi - \phi\|_{\mathcal{X}} = 0.$$

Étant donné un semi-groupe $(S_t)_{t \geq 0}$, on définit un unique opérateur non borné A de domaine

$$\mathcal{D}(A) := \{ \phi \in \mathcal{X} ; \lim_{t \searrow 0} \frac{1}{t} (S_t \phi - \phi) \text{ existe} \},$$

et pour $\phi \in \mathcal{D}(A)$,

$$A\phi = \lim_{t \searrow 0} \frac{1}{t} (S_t \phi - \phi).$$

Alors pour tout $\phi \in \mathcal{D}(A)$, l'application $u : t \mapsto S_t \phi$ est dérivable et vérifie $u' = Au$. Autrement dit, S_t joue le rôle de e^{tA} (défini seulement pour les opérateurs A continus). Plus généralement, si $u_0 \in \mathcal{D}(A)$ et $f \in \mathcal{C}(0, t; \mathcal{D}(A))$, la *formule de Duhamel*

$$u(t) = S_t u_0 + \int_0^t S_{t-\tau} f(\tau) d\tau.$$

résout le problème de Cauchy avec second membre :

$$u' = Au + f, \quad u(0) = u_0.$$

Pour $u_0 \in \mathcal{X}$ et $f \in L^1(0, t; \mathcal{X})$, la formule de Duhamel fournit seulement une « *mild solution* », terme que l'on pourrait traduire par « solution allégée » (et non « douce », car ce mot a un autre sens mathématique), mais qu'on laisse en anglais dans le texte en général. Une question naturelle mais difficile est de savoir, pour un opérateur non borné donné, s'il est le générateur infinitésimal d'un \mathcal{C}^0 -semi-groupe. Le théorème de Hille-Yosida (voir par exemple [3, ch.VII], ou [13]) répond à cette question pour certains types d'opérateurs. Un cas plus facile est celui des opérateurs *anti-autoadjoints* (voir l'appendice A pour la définition).

Groupes d'opérateurs unitaires On parle de groupe à un paramètre $(S_t)_{t \in \mathbb{R}}$ si pour tout $t \in \mathbb{R}$, S_t est un opérateur inversible sur un espace de Hilbert et $S_{-t} = S_t^{-1}$, et de groupe d'opérateurs unitaires si on a de plus $S_{-t} = S_t^*$ pour tout $t \in \mathbb{R}$.

Théorème I.2 (*Stone*)

Si A est un opérateur anti-autoadjoint à domaine dense dans un espace de Hilbert \mathcal{H} , c'est le générateur infinitésimal d'un groupe d'opérateurs unitaires.

Pour une démonstration, voir par exemple [13, p.41].

Les équations d'Airy et de Schrödinger sont précisément de la forme $u' = Au$, où A est un opérateur anti-autoadjoint à domaine dense dans L^2 . Notons au passage que toutes les équations de cette forme conservent la norme L^2 , puisque pour une solution (suffisamment régulière) on a

$$\frac{d}{dt} \int |u(t, \mathbf{x})|^2 d\mathbf{x} = 2\operatorname{Re} \langle u(t, \mathbf{x}), Au(t, \mathbf{x}) \rangle = 0.$$

Pour l'équation d'Airy, $\mathcal{D} = H^3$, et pour l'équation de Schrödinger, $\mathcal{D} = H^2$. Les groupes associés sont donnés à l'aide leur noyau de Green G_t par $S_t u_0 = u_0 * G_t$.

On cherche des estimations a priori de la forme

$$\|S_t u_0\|_{L^q(\mathbb{R}; L^r(\mathbb{R}^d))} \lesssim \|u_0\|_{L^2(\mathbb{R}^d)},$$

connaissant une estimation de la forme

$$(I.3) \quad \|S_t u_0\|_{L^r(\mathbb{R}^d)} \lesssim |t|^{-d\alpha(r)} \|u_0\|_{L^{r'}(\mathbb{R}^d)}.$$

Argument T^*T

L'idée (en s'inspirant de la formule de Duhamel) est de voir l'opérateur

$$\begin{array}{ccc} L^2(\mathbb{R}^d) & \rightarrow & L^\infty(\mathbb{R}; L^2(\mathbb{R}^d)) \\ v & \mapsto & S.v \end{array}$$

comme l'adjoint de l'opérateur

$$\begin{array}{ccc} T : L^1(\mathbb{R}; L^2(\mathbb{R}^d)) & \rightarrow & L^2(\mathbb{R}^d) \\ f & \mapsto & Tf := \int S_{-s} f(s) ds. \end{array}$$

En effet,

$$\langle T^* v, f \rangle = \langle v, Tf \rangle_{L^2(\mathbb{R}^d)} = \int \langle v, S_{-s} f(s) \rangle_{L^2(\mathbb{R}^d)} ds = \int \langle S_s v, f(s) \rangle_{L^2(\mathbb{R}^d)} ds.$$

Ainsi on a

$$\begin{array}{ccc} T^*T : L^1(\mathbb{R}; L^2(\mathbb{R}^d)) & \rightarrow & L^\infty(\mathbb{R}; L^2(\mathbb{R}^d)) \\ f & \mapsto & T^*Tf; \quad (T^*Tf)(\tau) = \int S_{\tau-s} f(s) ds. \end{array}$$

Le fait que T^*T soit continu provient en fait du lemme suivant, qui impliquera également, après application de l'inégalité de Hardy-Littlewood-Sobolev, l'estimation a priori cherchée.

Lemme I.1

Soit \mathcal{H} un espace de Hilbert, \mathcal{X} un espace de Banach, et \mathcal{D} un sous-espace vectoriel

dense dans \mathcal{X} . $T : \mathcal{D} \rightarrow \mathcal{H}$ opérateur linéaire $T^* : \mathcal{H} \rightarrow \mathcal{D}^*$ son adjoint algébrique ($\mathcal{X}^* \subset \mathcal{D}^*$ le dual algébrique de \mathcal{D})

$$\langle T^*v, f \rangle = \langle v, Tf \rangle, \quad \forall f \in \mathcal{D}, \forall v \in \mathcal{H}.$$

Les trois conditions suivantes sont équivalentes.

1. il existe $a \geq 0$ tel que, pour tout $f \in \mathcal{D}$,

$$\|Tf\|_{\mathcal{H}} \leq a\|f\|_{\mathcal{X}}.$$

2. l'image de T^* est incluse dans \mathcal{X}^* et il existe $b \geq 0$ tel que, pour tout $v \in \mathcal{H}$,

$$\|T^*v\|_{\mathcal{X}^*} \leq b\|v\|_{\mathcal{H}}.$$

3. l'image de T^*T est incluse dans \mathcal{X}^* et il existe $c \geq 0$ tel que, pour tout $f \in \mathcal{D}$,

$$\|T^*Tf\|_{\mathcal{X}^*} \leq c\|f\|_{\mathcal{X}}.$$

Démonstration : $1 \implies 2$: pour $v \in \mathcal{H}$ et $f \in \mathcal{D}$,

$$|\langle T^*v, f \rangle| \leq \|v\|_{\mathcal{H}} \|Tf\|_{\mathcal{H}} \leq a \|v\|_{\mathcal{H}} \|f\|_{\mathcal{X}}.$$

Donc d'après le théorème de Hahn-Banach (voir l'appendice A), T^*v se prolonge en une forme linéaire continue sur \mathcal{X} et

$$\|T^*v\|_{\mathcal{X}^*} \leq a\|v\|_{\mathcal{H}}.$$

$2 \implies 1$: pour $v \in \mathcal{H}$ et $f \in \mathcal{D}$,

$$|\langle v, Tf \rangle| \leq \|T^*v\|_{\mathcal{X}^*} \|f\|_{\mathcal{X}} \leq b \|v\|_{\mathcal{H}} \|f\|_{\mathcal{X}},$$

donc

$$\|Tf\|_{\mathcal{H}} \leq b\|f\|_{\mathcal{X}}.$$

Par ailleurs, 1 et 2 impliquent trivialement 3.

$3 \implies 1$: pour $f \in \mathcal{D}$,

$$|\langle Tf, Tf \rangle| \leq \|T^*Tf\|_{\mathcal{X}^*} \|f\|_{\mathcal{X}} \leq c\|f\|_{\mathcal{X}}^2.$$

■

Pour conclure, on va avoir recours au résultat (difficile) suivant.

Théorème I.3 (Hardy-Littlewood-Sobolev)

Soient $a, p, q \in]1, +\infty[$, avec $p < q$ et

$$\frac{1}{a} + \frac{1}{p} = 1 + \frac{1}{q}.$$

Soit

$$\phi_a : \mathbf{y} \mapsto \|\mathbf{y}\|^{-n/a}.$$

Alors $u \mapsto u * \phi_a$ est continue de $L^p(\mathbb{R}^n)$ dans $L^q(\mathbb{R}^n)$.

Remarque I.2

Si on avait $\phi_a \in L^a$, ce serait un résultat classique de convolution (qui autoriserait même les cas extrêmes $a = 1$, $p = 1$, $p = q$, $q = +\infty$). Mais ce n'est précisément pas le cas : la fonction ϕ_a n'appartient à aucun espace L^k . Malgré cela, ce théorème montre qu'on peut la convoler avec $u \in L^p$ pour $p > 1$ et obtenir une fonction dans L^q avec $p < q < +\infty$.

Voir [8, p. 117 et suivantes] pour une démonstration.

L'inégalité de Hardy-Littlewood-Sobolev (avec $n = 1$) et l'inégalité (I.3) impliquent que l'opérateur

$$T^*T : f \mapsto T^*Tf; \quad (T^*Tf)(\tau) = \int S_{\tau-s}f(s) \, ds$$

est continu de $L^{q'}(\mathbb{R}; L^{r'}(\mathbb{R}^d))$ dans $L^q(\mathbb{R}; L^r(\mathbb{R}^d))$ pour $0 < 2/q = d\alpha(r) < 1$. En effet, pour tout $\phi \in L^{q'}(\mathbb{R}; L^{r'}(\mathbb{R}^d))$, on a par l'inégalité de Hölder en espace

$$\begin{aligned} |\langle \phi, T^*Tf \rangle| &\leq \iint \|G_{\tau-s} * f(s)\|_{L^r(\mathbb{R}^d)} \|\phi(\tau)\|_{L^{r'}(\mathbb{R}^d)} \, ds d\tau \\ &\lesssim \iint |\tau - s|^{-d\alpha(r)} \|f(s)\|_{L^{r'}(\mathbb{R}^d)} \|\phi(\tau)\|_{L^{r'}(\mathbb{R}^d)} \, ds d\tau \end{aligned}$$

d'après (I.3), d'où par le théorème de Fubini et l'inégalité de Hölder en temps,

$$|\langle \phi, T^*Tf \rangle| \lesssim \left\| \tau \mapsto \int |\tau - s|^{-d\alpha(r)} \|f(s)\|_{L^{r'}(\mathbb{R}^d)} \, ds \right\|_{L^q(\mathbb{R})} \|\phi\|_{L^{q'}(\mathbb{R}; L^{r'}(\mathbb{R}^d))}.$$

Et d'après l'inégalité de Hardy-Littlewood-Sobolev,

$$\left\| \tau \mapsto \int |\tau - s|^{-d\alpha(r)} \|f(s)\|_{L^{r'}(\mathbb{R}^d)} \, ds \right\|_{L^q(\mathbb{R})} \lesssim \|f\|_{L^{q'}(\mathbb{R}; L^{r'}(\mathbb{R}^d))}$$

pour $d\alpha(r) + \frac{1}{q'} = 1 + \frac{1}{q}$, c'est-à-dire $d\alpha(r) = 2/q$, sous les conditions $0 < d\alpha(r) < 1$, $1 < q' < q < +\infty$.

D'après le lemme I.1, ceci implique que T^* est continu de L^2 dans $L^q(\mathbb{R}; L^r(\mathbb{R}^d))$.

Théorème I.4

Soit (S_t) un groupe d'opérateurs unitaires sur $L^2(\mathbb{R}^d)$ vérifiant l'inégalité de dispersion (I.3) pour $r \in [2, +\infty[$. Alors il vérifie l'inégalité de Strichartz

$$\|S \cdot u_0\|_{L^q(\mathbb{R}; L^r(\mathbb{R}^d))} \lesssim \|u_0\|_{L^2(\mathbb{R}^d)},$$

pour r et q tels que $0 < 2/q = d\alpha(r) < 1$.

Certaines valeurs extrêmes des paramètres sont en fait autorisées. Pour l'équation de Schrödinger en particulier, où $\alpha(r) = 1/2 - 1/r$, tous les triplets (q, r, d) pour lesquels $2/q = d/2 - d/r$ et $2 \leq q, r \leq +\infty$ sont autorisés, sauf $(q, r, d) = (2, +\infty, 2)$ (voir [14, p. 74]).

Équation des ondes

Rappelons que la solution du problème de Cauchy

$$\partial_{tt}^2 u - c^2 \Delta u = 0, \quad u|_{t=0} = u_0, \quad \partial_t u|_{t=0} = u_1,$$

s'exprime à l'aide de deux noyaux K_t^0 et K_t^1 qui sont des distributions :

$$u(t, \cdot) = u_0 * K_t^0 + u_1 * K_t^1,$$

$$K_t^0 = \partial_t K_t^1 = \mathcal{F}^{-1}(\xi \mapsto \cos(ct\|\xi\|)), \quad K_t^1 = \mathcal{F}^{-1}\left(\xi \mapsto \frac{\sin(ct\|\xi\|)}{c\|\xi\|}\right).$$

Nous allons ici nous concentrer sur la dimension $d = 3$. Il n'est dans ce cas pas très difficile de calculer K_t^1 (ce qui permet entre autres de retrouver la formule de Kirchhoff). On trouve :

$$K_t^1 = \frac{1}{4\pi c^2 t} \sigma_{ct},$$

où σ_{ct} désigne la distribution supportée par la sphère de rayon ct , définie simplement par

$$\langle \sigma_r, \phi \rangle = \int_{\|\mathbf{x}\|=r} \phi(\mathbf{x}) d\sigma(\mathbf{x}) = \int_0^{2\pi} \int_0^\pi \phi(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta d\theta d\varphi.$$

En effet, comme σ_r est à support compact, sa transformée de Fourier est la fonction définie par

$$\widehat{\sigma}_r(\xi) = \langle \sigma_r, e^{-i\xi \cdot \mathbf{x}} \rangle,$$

et comme σ_r est à symétrie radiale, $\widehat{\sigma}_r$ aussi. Il suffit donc de calculer

$$\widehat{\sigma}_r(0, 0, \xi_3) = \int_0^{2\pi} \int_0^\pi e^{-i\xi_3 r \cos \theta} r^2 \sin \theta d\theta d\varphi = 4\pi r \frac{\sin(\xi_3 r)}{\xi_3},$$

d'où l'on déduit,

$$\widehat{\sigma}_r(\xi) = 4\pi r \frac{\sin(r\|\xi\|)}{\|\xi\|}.$$

Remarque I.3

Attention, les opérateurs de convolution Q_t et P_t associés respectivement aux noyaux K_t^0 et K_t^1 ne sont pas des semi-groupes ! D'autre part, comme ce sont des multiplicateurs de Fourier à symbole réel ($\mathcal{F}(K_t^0)$ et $\mathcal{F}(K_t^1)$ sont à valeurs réelles), ce sont des opérateurs auto-adjoints.

On peut néanmoins adapter la méthode utilisée plus haut pour les groupes d'opérateurs unitaires et montrer des inégalités de Strichartz pour l'équation des ondes. Ces inégalités s'écrivent de façon générale :

$$(I.4) \quad \begin{aligned} & \|u\|_{L^p(\mathbb{R}^+; L^q(\mathbb{R}^d))} \lesssim \|(u_1, \nabla u_0)\|_{L^2(\mathbb{R}^d)}, \quad \text{pour } \frac{1}{p} + \frac{3}{q} = \frac{d}{2} - 1, \\ & \frac{2}{p} + \frac{d-1}{q} \leq \frac{d-1}{2}, \quad 2 \leq p \leq q \leq +\infty, \quad d \geq 2, \quad (p, q, d) \neq (2, +\infty, 3). \end{aligned}$$

Nous allons nous contenter de traiter, en dimension $d = 3$, le cas de P_t , ce qui revient à supposer $u_0 \equiv 0$. L'idée est à nouveau de voir l'opérateur

$$\begin{aligned} L^2(\mathbb{R}^d) &\rightarrow L^\infty(\mathbb{R}; L^2(\mathbb{R}^d)) \\ v &\mapsto P.v \end{aligned}$$

comme l'adjoint d'un opérateur T pour lequel on saura estimer T^*T . En l'occurrence, on considère

$$\begin{aligned} T : L^1(\mathbb{R}; L^2(\mathbb{R}^d)) &\rightarrow L^2(\mathbb{R}^d) \\ f &\mapsto Tf := \int P_t f(t) dt. \end{aligned}$$

On a bien $T^*v = P.v$ car

$$\langle T^*v, f \rangle = \langle v, Tf \rangle_{L^2(\mathbb{R}^d)} = \int \langle v, P_t f(t) \rangle_{L^2(\mathbb{R}^d)} dt = \int \langle P_t v, f(s) \rangle_{L^2(\mathbb{R}^d)} ds.$$

Et par ailleurs,

$$\begin{aligned} T^*T : L^1(\mathbb{R}; L^2(\mathbb{R}^d)) &\rightarrow L^\infty(\mathbb{R}; L^2(\mathbb{R}^d)) \\ f &\mapsto T^*Tf; \quad (T^*Tf)(\tau) = \int P_\tau P_s f(s) ds. \end{aligned}$$

L'opérateur composé $P_\tau P_s$ est encore un multiplicateur de Fourier, de symbole (réel)

$$\frac{\sin(c\tau\|\xi\|) \sin(cs\|\xi\|)}{c^2\|\xi\|^2} = \frac{1}{2} (\widehat{K_{\tau-s}}(\xi) - \widehat{K_{\tau+s}}(\xi)),$$

où

$$\widehat{K_t}(\xi) := \frac{\cos(ct\|\xi\|)}{c^2\|\xi\|^2}.$$

Ce nouveau noyau K_t a l'avantage par rapport à K_t^1 d'être une fonction ! On démontre en effet (voir l'appendice B) que pour tout $t \geq 0$,

$$K_t(\mathbf{x}) = \frac{H(\|\mathbf{x}\| - ct)}{4\pi\|\mathbf{x}\|},$$

où H est la *fonction de Heaviside* (valant 1 sur \mathbb{R}^{+*} et 0 sur \mathbb{R}^-). De plus ce noyau vérifie une inégalité de dispersion : on vérifie par le calcul (en coordonnées sphériques) que pour tout $r > 3$,

$$\|K_t\|_{L^r(\mathbb{R}^3)} \lesssim t^{-1+3/r}.$$

Par suite, pour $\tau, s \geq 0$, l'opérateur $P_\tau P_s$ (de convolution par $(K_{\tau-s} - K_{\tau+s})/2$) vérifie, par une inégalité de convolution standard,

$$\|P_\tau P_s f\|_{L^q(\mathbb{R}^3)} \lesssim (|\tau - s|^{-1+3/r} + |\tau + s|^{-1+3/r}) \|f\|_{L^{q'}(\mathbb{R}^3)}, \quad \text{pour } \frac{1}{r} + \frac{1}{q'} = 1 + \frac{1}{q},$$

c'est-à-dire pour $q = 2r$. Par l'inégalité de Hardy-Littlewood-Sobolev on obtient alors

$$\|T^*Tf\|_{L^p(\mathbb{R}^+; L^q(\mathbb{R}^3))} \lesssim \|f\|_{L^{p'}(\mathbb{R}^+; L^{q'}(\mathbb{R}^3))}$$

pour $2/p = 1 - 3/r$, c'est-à-dire $1/p + 3/q = 1/2$. Du lemme I.1 on déduit donc

$$\|P.v\|_{L^p(\mathbb{R}^+; L^q(\mathbb{R}^3))} \lesssim \|v\|_{L^2(\mathbb{R}^3)},$$

ce qui revient à (I.4) dans le cas particulier $u_0 \equiv 0$.

Remarque I.4

Si l'on considère l'équation des ondes comme le système d'ordre 1

$$\partial_t U = \begin{pmatrix} 0 & c^2 \operatorname{div} \\ \nabla & 0 \end{pmatrix} U,$$

la conservation de l'énergie s'interprète comme la conservation de la norme L^2 de U . Plus généralement, toutes les normes H^s sont conservées : pour tout $s \in \mathbb{R}$, pour tout $t \geq 0$,

$$\|U(t)\|_{H^s(\mathbb{R}^d)} = \|U(0)\|_{H^s(\mathbb{R}^d)} = \|(u_1, \nabla u_0)\|_{H^s(\mathbb{R}^d)}.$$

En particulier,

$$\|u(t)\|_{\dot{H}^s(\mathbb{R}^d)} \leq \|\nabla u(t)\|_{H^{s-1}(\mathbb{R}^d)} \lesssim \|u_1\|_{H^{s-1}(\mathbb{R}^d)} + \|\nabla u_0\|_{H^{s-1}(\mathbb{R}^d)},$$

et comme $\dot{H}^s(\mathbb{R}^d)$ s'injecte continûment dans $L^q(\mathbb{R}^d)$ pour $s \in]0, d/2[$ et $1/q = 1/2 - s/d$, on en déduit l'estimation

$$\|u\|_{L^\infty(\mathbb{R}^+; L^q(\mathbb{R}^d))} \lesssim \|u_1\|_{H^{s-1}(\mathbb{R}^d)} + \|\nabla u_0\|_{H^{s-1}(\mathbb{R}^d)}.$$

On trouve encore beaucoup d'autres estimations de type Strichartz dans la littérature. Citons par exemple, d'après les travaux de Koenig, Ponce et Vega (voir [14, pp. 205-206]), les inégalités suivantes pour l'équation d'Airy (avec $a = 0$) :

$$\|u\|_{L^2(0,T; L^\infty(\mathbb{R}))} \lesssim \|u_0\|_{H^s}, \quad s > 3/4,$$

$$\|u\|_{L^4(0,T; L^\infty(\mathbb{R}))} \lesssim \|u_0\|_{H^s}, \quad s \geq 1/4,$$

$$\|\partial_x^{s+1} u\|_{L^\infty(\mathbb{R}; L^2(0,T))} \lesssim \|u_0\|_{H^s}, \quad s \geq 0.$$

Chapitre II

EDP dispersives non-linéaires

1 Exemples et vocabulaire

On va s'intéresser dans ce chapitre à des versions non-linéaires des EDP rencontrées précédemment.

1.1 NLS et NLW

Les EDP *semi-linéaires* sont celles où la non-linéarité est d'ordre 0. C'est le cas en particulier pour l'équation de Schrödinger non-linéaire (connue sous le sigle anglais NLS)

$$i\partial_t u + \Delta u = f(u),$$

et l'équation des ondes non-linéaire (connue sous le sigle anglais NLW)

$$\partial_{tt}^2 u - c^2 \Delta u + f(u) = 0.$$

L'équation des ondes non-linéaire avec $f(u) = \sin u$ est appelée *équation de sine-Gordon* (par analogie phonétique avec l'équation de Klein-Gordon, qui en est la version linéarisée au voisinage de 0). Un autre type très étudié de non-linéarité est

$$f(u) = \mu |u|^{p-1} u,$$

qui « dérive du potentiel »

$$V(u) = \frac{\mu}{p+1} |u|^{p+1}$$

(au sens où $\nabla_u V = f(u)$ si u est complexe, ou simplement $V'(u) = f(u)$ si u est réel). En particulier, si p est un entier impair (cas dit « algébrique »), l'application $u \in \mathbb{C} \mapsto \mu |u|^{p-1} u$ est de classe \mathcal{C}^∞ . Selon le signe de μ on parle de non-linéarité (ou d'équation) *focalisante* ou *défocalisante*. Ceci s'interprète notamment en terme de solutions explicites obtenues par *séparation de variables*. En effet, si l'on cherche des solutions sous la forme

$$u(t, \mathbf{x}) = e^{i\boldsymbol{\xi} \cdot \mathbf{x}} v(t)$$

avec $|v(t)| = \text{constante}$ (ce qui généralise les OPPM), on trouve

$$v(t) = \alpha e^{i(\|\boldsymbol{\xi}\|^2 + \mu |\alpha|^{p-1})t} \quad \text{pour (NLS),}$$

$$v(t) = \alpha e^{\pm it \sqrt{c^2 \|\xi\|^2 + \mu |\alpha|^{p-1}}} \quad \text{pour (NLW)}.$$

Dans ce deuxième cas il faut en outre supposer $c^2 \|\xi\|^2 + \mu |\alpha|^{p-1} \geq 0$, ce qui est automatique pour $\mu \geq 0$ mais n'est valable que pour les « hautes » fréquences ($\|\xi\| \geq \sqrt{-\mu} |\alpha|^{(p-1)/2}/c$) si $\mu < 0$. On constate sur ces formules que $\mu > 0$ ajoute une contribution « positive » à la dispersion linéaire, tandis que $\mu < 0$ va à son encontre : le premier cas est dit défocalisant et le second est focalisant.

On observe par ailleurs que (NLS) et (NLW) sont invariantes par certains changements d'échelles. Plus précisément, (NLS) est invariante par

$$u \mapsto \tilde{u}; \quad \tilde{u}(\tilde{t}, \tilde{x}) = \lambda^{-2/(p-1)} u(\tilde{t}/\lambda^2, \tilde{x}/\lambda),$$

et (NLW) est invariante par

$$u \mapsto \tilde{u}; \quad \tilde{u}(\tilde{t}, \tilde{x}) = \lambda^{-2/(p-1)} u(\tilde{t}/\lambda, \tilde{x}/\lambda),$$

pour tout $\lambda > 0$. Ceci conduit à définir l'*indice de régularité critique*

$$s_c = \frac{d}{2} - \frac{2}{p-1},$$

de sorte que la norme $\dot{H}^{s_c}(\mathbb{R}^d)$ soit invariante par

$$u_0 \mapsto \tilde{u}_0; \quad \tilde{u}_0(\tilde{x}) = \lambda^{-2/(p-1)} u_0(\tilde{x}/\lambda),$$

et la norme de $\dot{H}^{s_c}(\mathbb{R}^d) \times \dot{H}^{s_c-1}(\mathbb{R}^d)$ soit invariante par

$$(u_0, u_1) \mapsto (\tilde{u}_0, \tilde{u}_1); \quad \tilde{u}_0(\tilde{x}) = \lambda^{-2/(p-1)} u_0(\tilde{x}/\lambda), \quad \tilde{u}_1(\tilde{x}) = \lambda^{-2/(p-1)-1} u_1(\tilde{x}/\lambda).$$

Les espaces de Sobolev d'indice $s > s_c$ seront alors dit *sous-critiques* et ceux d'indice $s < s_c$ seront dits *sur-critiques*. (L'inversion entre les inégalités et les préfixes sous/sur est motivée par le fait que le problème de Cauchy sera moins difficile à traiter dans les espaces d'indice élevé.)

Les deux équations, (NLS) et (NLW), se mettent sous la forme semi-linéaire abstraite

$$(II.1) \quad \partial_t U = AU + N(U),$$

où A est un opérateur différentiel à coefficients constants agissant sur des fonctions à valeurs dans \mathbb{R}^n (\mathbb{R}^2 identifié avec \mathbb{C} pour (NLS), et \mathbb{R}^{d+2} pour (NLW)), et $N(U) = F \circ U$ avec $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$. On supposera dans la suite que cette fonction F est de classe \mathcal{C}^∞ et vérifie $F(0) = 0$. Concrètement, pour l'équation de Schrödinger,

$$U = \begin{pmatrix} \operatorname{Re} u \\ \operatorname{Im} u \end{pmatrix}, \quad A = \begin{pmatrix} 0 & -\Delta \\ \Delta & 0 \end{pmatrix}, \quad F(U) = \begin{pmatrix} \operatorname{Im} f(u) \\ \operatorname{Re} f(u) \end{pmatrix},$$

et pour l'équation des ondes,

$$U = \begin{pmatrix} u \\ v = \frac{1}{c} \partial_t u \\ w = \nabla u \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & c \operatorname{div} \\ 0 & c \nabla & 0 \end{pmatrix}, \quad F(U) = \begin{pmatrix} cv \\ -\frac{1}{c} f(u) \\ 0 \end{pmatrix}.$$

On rappelle que par le théorème de Stone, comme les opérateurs A ci-dessus sont anti-autoadjoints, ils sont générateurs de groupes d'opérateurs unitaires.

1.2 KdV et Boussinesq

Dans les EDP *quasi-linéaires*, la non-linéarité est en fait linéaire par rapport à la dérivée d'ordre le plus élevé. C'est le cas en particulier pour l'équation de Korteweg-de Vries (généralisée)

$$\partial_t u + \partial_{xxx}^3 u = \partial_x f(u),$$

et l'équation de Boussinesq généralisée

$$\partial_{tt}^2 u - \partial_{xx}^2 f(u) + \partial_{xxx}^3 u = 0.$$

Cette dernière peut être vue formellement (après intégration en x) comme le système d'ordre 3, que l'on appellera *système de Boussinesq* :

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \partial_x \begin{pmatrix} -v \\ -f(u) + \partial_{xx}^2 u \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dans l'équation de Korteweg-de Vries classique (KdV), la non-linéarité est quadratique : $f'(u) = u$ à un facteur multiplicatif près. Dans ce cas on a une invariance par le changement d'échelle

$$u \mapsto \tilde{u}; \quad \tilde{u}(\tilde{t}, \tilde{x}) = \lambda^{-2} u(\tilde{t}/\lambda^3, \tilde{x}/\lambda),$$

et l'indice critique est $s_c = -3/2$. De même, pour une non-linéarité quadratique, le système de Boussinesq est invariant par le changement d'échelle

$$(u, v) \mapsto (\tilde{u}, \tilde{v}); \quad \tilde{u}(\tilde{t}, \tilde{x}) = \lambda^{-2} u(\tilde{t}/\lambda^2, \tilde{x}/\lambda), \quad \tilde{v}(\tilde{t}, \tilde{x}) = \lambda^{-3} v(\tilde{t}/\lambda^2, \tilde{x}/\lambda),$$

et l'espace critique est $\dot{H}^{-3/2}(\mathbb{R}) \times \dot{H}^{-5/2}(\mathbb{R})$.

L'équation de KdV et le système de Boussinesq se mettent sous la forme quasi-linéaire abstraite

$$(II.2) \quad \partial_t U + \partial_x F(U) + K \partial_{xxx}^3 U = 0,$$

où l'inconnue U est scalaire (pour KdV, avec $K = 1$) ou à valeurs dans \mathbb{R}^2 (pour le système de Boussinesq, avec $K = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$).

2 Problème de Cauchy

Au chapitre précédent, le problème de Cauchy a été vu essentiellement au travers de formules de résolution (noyaux de Green) et estimations *a priori*. Il est temps de définir précisément la notion de solution, et en fait les différentes notions de solutions, et celle de problème de Cauchy bien posé, pour des équations de la forme (II.1) ou (II.2).

2.1 Les différentes notions de solutions

La notion la plus facile à appréhender est celle de *solution classique* : une fonction $(t, \mathbf{x}) \mapsto U(t, \mathbf{x})$ est solution classique de (II.1), respectivement (II.2), si elle est de classe \mathcal{C}^1 en t et de classe \mathcal{C}^m en \mathbf{x} , où m est l'ordre de l'opérateur différentiel A , respectivement $m = 3$. Pour une solution classique U , une *donnée initiale* est une fonction U_0 de classe

\mathcal{C}^m telle que $U(0, \cdot) = U_0$. À l'autre extrémité, une *solution faible* (ou solution au sens des distributions) du problème de Cauchy sur $[0, T]$ pour (II.1), respectivement (II.2), avec donnée initiale $U_0 \in L_{\text{loc}}^\infty(\mathbb{R}^d)$ est une fonction $U \in L_{\text{loc}}^\infty([0, T] \times \mathbb{R}^d; \mathbb{R}^n)$ vérifiant pour tout $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d; (\mathbb{R}^n)')$,

$$\int_0^T \int_{\mathbb{R}^d} \langle (\partial_t \varphi - A^* \varphi)(t, \mathbf{x}), U(t, \mathbf{x}) \rangle + \langle \varphi(t, \mathbf{x}), F(U(t, \mathbf{x})) \rangle d\mathbf{x} dt + \int_0^T \langle U_0(\mathbf{x}), \varphi(0, \mathbf{x}) \rangle d\mathbf{x} = 0,$$

respectivement,

$$\int_0^T \int_{\mathbb{R}^d} \langle (\partial_t \varphi + K^t \partial_{xxx}^3 \varphi)(t, \mathbf{x}), U(t, \mathbf{x}) \rangle + \langle \partial_x \varphi(t, \mathbf{x}), F(U(t, \mathbf{x})) \rangle d\mathbf{x} dt + \int_0^T \langle U_0(\mathbf{x}), \varphi(0, \mathbf{x}) \rangle d\mathbf{x} = 0.$$

Définition II.1

Étant donné un espace fonctionnel \mathcal{H} de données initiales, on appelle

- solution faible dans \mathcal{H} sur $[0, T]$ une solution faible appartenant en outre à l'espace $L^\infty(0, T; \mathcal{H})$,
- solution forte dans \mathcal{H} sur $[0, T]$ une solution faible appartenant en outre à l'espace $\mathcal{C}(0, T; \mathcal{H})$.

Remarque II.1

Si A est un opérateur borné sur $\mathcal{H} = L^2(\mathbb{R}^d; \mathbb{R}^n)$ et si $N : U \in \mathcal{H} \mapsto F \circ U \in \mathcal{H}$, alors (II.1) est en fait un équation différentielle ordinaire dans \mathcal{H} , et les trois notions, solution classique, solution forte et solution faible, coïncident pour l'équation (II.1).

En effet, une solution classique est toujours une solution faible (ceci se voit juste en intégrant la relation $\langle \varphi, \partial_t U - AU - F \circ U \rangle = 0$), et même si l'expression n'est pas très heureuse, une solution forte est solution faible, par définition. Le « pivot » de la réciproque est la formulation intégrale de (II.1), à savoir

$$(II.3) \quad U(t) = U(0) + \int_0^t (AU(\tau) + F \circ U(\tau)) d\tau.$$

Si U est solution faible dans \mathcal{H} sur $[0, T]$, on a par définition $U \in L^\infty(0, T; \mathcal{H})$ et donc $F \circ U \in L^\infty(0, T; \mathcal{H})$ aussi. Or par définition d'une solution faible on a pour tout $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d; (\mathbb{R}^n)')$,

$$\int_0^T \int_{\mathbb{R}^d} \langle \partial_t \varphi(t, \mathbf{x}), U(t, \mathbf{x}) \rangle + \langle \varphi(t, \mathbf{x}), AU(t, \mathbf{x}) + F(U(t, \mathbf{x})) \rangle d\mathbf{x} dt = 0.$$

Comme $AU + F \circ U \in L^\infty(0, T; L^2(\mathbb{R}^d; \mathbb{R}^n))$, ceci montre que $U \in W^{1,1}(0, T; L^2(\mathbb{R}^d; \mathbb{R}^n))$, qui s'injecte dans $\mathcal{C}(0, T; L^2(\mathbb{R}^d; \mathbb{R}^n))$, et que U satisfait la formule intégrale (II.3). Cette même formule montre ensuite que $U \in \mathcal{C}^1(0, T; L^2(\mathbb{R}^d; \mathbb{R}^n))$, et en la dérivant on obtient que U est solution classique.

Remarque II.2

Bien entendu, pour les cas qui nous intéressent, A n'est pas un opérateur borné sur L^2 , puisque c'est un opérateur différentiel, et F est une « vraie » non-linéarité, qui ne

préserve pas L^2 .

Définition II.2

Le problème de Cauchy est localement bien posé dans un espace fonctionnel \mathcal{H} si pour tout $U_* \in \mathcal{H}$ il existe $T > 0$, un voisinage \mathcal{V} de U_* dans \mathcal{H} , et un sous-ensemble \mathcal{E} de $\mathcal{C}(0, T; \mathcal{H})$ tels que,

1. pour tout $U_0 \in \mathcal{V}$, il existe un unique $U \in \mathcal{E}$ qui soit solution forte dans \mathcal{H} sur $[0, T]$ pour la donnée initiale U_0
2. l'application

$$\begin{aligned} \mathcal{V} &\rightarrow \mathcal{C}(0, T; \mathcal{H}) \\ U_0 &\mapsto U \end{aligned}$$

est continue.

Si c'est vrai avec $\mathcal{E} = \mathcal{C}(0, T; \mathcal{H})$ on dit que le problème de Cauchy est inconditionnellement (localement) bien posé. Si c'est vrai pour T arbitrairement grand on dit que le problème de Cauchy est globalement bien posé.

Les résultats du chapitre I montrent que les problèmes de Cauchy linéaires (c'est-à-dire avec $F \equiv 0$) sont globalement bien posés dans $H^s(\mathbb{R}^d)$ pour l'équation de Schrödinger et l'équation d'Airy, et dans $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$ pour l'équation des ondes vue comme un système d'ordre 1. Qu'en est-il pour les problèmes non-linéaires ?

2.2 Méthode des semi-groupes

Considérons une EPD de la forme (II.1), où A est un opérateur non borné, anti-autoadjoint sur un espace de Hilbert \mathcal{H} . En considérant $V(t) := S_{-t}U(t)$, où $(S_t)_{t \in \mathbb{R}}$ est le groupe d'opérateurs unitaires engendré par A , et en remarquant que

$$\langle \partial_t \varphi + A^* \varphi, U \rangle_{\mathcal{H}} = \langle \partial_t (S_{-t} \varphi(t)), V \rangle_{\mathcal{H}},$$

on obtient en procédant comme pour (II.3) une formulation intégrale pour les solutions fortes :

$$(II.4) \quad U(t) = S_t U_0 + \int_0^t S_{t-\tau} (F \circ U)(\tau) d\tau,$$

pour autant que $t \mapsto F \circ U(t)$ appartienne à $L^1(0, T; \mathcal{H})$. Lorsque $\mathcal{H} = H^s(\mathbb{R}^d)$ par exemple, ceci loin d'être automatiquement le cas. Cependant, on a quelques résultats fort utiles concernant ces espaces H^s et les non-linéarités.

Théorème II.1 (estimations de Moser)

- Si $u, v \in H^s(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$ avec $s \geq 0$, on a $uv \in H^s(\mathbb{R}^d)$ avec

$$\|uv\|_{H^s(\mathbb{R}^d)} \lesssim \|u\|_{L^\infty(\mathbb{R}^d)} \|v\|_{H^s(\mathbb{R}^d)} + \|v\|_{L^\infty(\mathbb{R}^d)} \|u\|_{H^s(\mathbb{R}^d)}.$$

- Si $f \in \mathcal{C}^\infty(\mathbb{R}; \mathbb{R})$ est telle que $f(0) = 0$, pour tout $s > d/2$, il existe une fonction continue $C : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ telle que pour tout $u \in H^s(\mathbb{R}^d)$ on a $f \circ u \in H^s(\mathbb{R}^d)$ avec

$$\|f \circ u\|_{H^s(\mathbb{R}^d)} \leq C(\|u\|_{L^\infty}) \|u\|_{H^s}.$$

- Pour tout $s > d/2$, il existe une fonction continue $C : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ telle que pour tous $u, v \in H^s(\mathbb{R}^d)$ vérifiant $\|(u, v)\|_{H^s(\mathbb{R}^d)} \leq \mu$ et $\|(u, v)\|_{L^\infty(\mathbb{R}^d)} \leq R$, on a

$$\|f \circ u - f \circ v\|_{H^s(\mathbb{R}^d)} \leq C(R) (\|u - v\|_{H^s(\mathbb{R}^d)} + \mu \|u - v\|_{L^\infty(\mathbb{R}^d)}).$$

Remarque II.3

D'après l'injection de Sobolev $H^s(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$, le premier point montre en particulier que $H^s(\mathbb{R}^d)$ est une algèbre de Banach lorsque $s > d/2$.

Théorème II.2

Pour une EPD de la forme (II.1) avec A un opérateur anti-autoadjoint sur $H^s(\mathbb{R}^d; \mathbb{R}^n)$, $s > d/2$, et $F \in \mathcal{C}^\infty(\mathbb{R}^n; \mathbb{R}^n)$ telle que $F(0) = 0$, le problème de Cauchy est localement inconditionnellement bien posé dans $H^s(\mathbb{R}^d; \mathbb{R}^n)$.

Démonstration : Le schéma de la démonstration est le même que pour les équations différentielles ordinaires, en utilisant la formule intégrale (II.4) (Duhamel) au lieu de la formule intégrale (II.3), qui est inutilisable lorsque A est un opérateur non borné. En fait ce schéma fonctionne dès que A est le générateur infinitésimal d'un semi-groupe de contractions $(S_t)_{t \geq 0}$, c'est-à-dire que tous les opérateurs S_t pour $t \geq 0$ sont de norme au plus 1.

On fixe une fois pour toutes $s > d/2$ et $U_* \in H^s(\mathbb{R}^d)$. On notera $\mu_* = \|U_*\|_{H^s}$ et $R_* = \|U_*\|_{L^\infty}$. Soient $\mu > 0$, $R > 0$,

$$U_0 \in \mathcal{V} := \{V \in H^s(\mathbb{R}^d; \mathbb{R}^n); \|V - U_*\|_{H^s} \leq \mu/8, \|V - U_*\|_{L^\infty} \leq R/2\},$$

$$\mathcal{E} := \{U \in \mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n)); \forall t \in [0, T], \|U(t) - U_*\|_{H^s} \leq \mu, \|U(t) - U_*\|_{L^\infty} \leq R\}.$$

On va construire une solution de $U \in \mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$ de (II.4) sur un intervalle $[0, T]$ assez petit de sorte que

$$\|U(t) - U_0\|_{H^s} \leq \mu/2, \quad \|U(t) - U_0\|_{L^\infty} \leq R/2,$$

ce qui impliquera évidemment $U \in \mathcal{E}$. Pour cela on met en place un schéma itératif. On pose $U^0 := U_0$. En supposant $U^k \in \mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$ et vérifiant les inégalités ci-dessus, on peut définir

$$U^{k+1}(t) := S_t U_0 + \int_0^t S_{t-\tau} (F \circ U^k)(\tau) d\tau,$$

et

$$\begin{aligned} \|U^{k+1}(t) - U_0\|_{H^s} &\leq \|S_t U_* - U_*\|_{H^s} + \|S_t(U_0 - U_*)\|_{H^s} + \|U_0 - U_*\|_{H^s} + \int_0^t \|F \circ U^k(\tau)\|_{H^s} d\tau. \\ &\leq \|S_t U_* - U_*\|_{H^s} + \mu/4 + TC(R_* + R)(\mu_* + 5\mu/8) \end{aligned}$$

d'après le théorème II.1. En choisissant $T \leq \mu / ((8\mu_* + 5\mu)C(R_* + R))$ tel que $\|S_t U_* - U_*\|_{H^s} \leq \mu/8$, on aura bien $\|U^{k+1}(t) - U_0\|_{H^s} \leq \mu/2$. Et si μ a été choisi au départ tel que $C_{d,s}\mu \leq R$, où $C_{d,s}$ désigne la norme de l'injection de Sobolev $H^s(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d)$, on aura évidemment aussi $\|U^{k+1}(t) - U_0\|_{L^\infty} \leq R/2$. Par construction, $U^{k+1} \in \mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$, puisque $t \mapsto S_t V$ est continue et $V \mapsto F \circ V$ est continue (par le

théorème II.1). L'étape suivante consiste à montrer que la suite $(U^k)_{k \in \mathbb{N}}$ est de Cauchy dans $\mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$. On a, pour $k \geq 1$,

$$U^{k+1}(t) - U^k(t) = \int_0^t S_{t-\tau} (F \circ U^k(\tau) - F \circ U^{k-1}(\tau)) d\tau,$$

et donc

$$\|U^{k+1}(t) - U^k(t)\|_{H^s} \leq L \int_0^t \|U^k(\tau) - U^{k-1}(\tau)\|_{H^s} d\tau,$$

avec $L := C(R_* + R)(1 + C_{d,s}(\mu_* + \mu))$. Ceci montre (via une récurrence facile) que

$$\|U^k(t) - U^{k-1}(t)\|_{H^s} \leq \frac{L^n t^n}{n!} \max_{t \in [0, T]} \|U^1(t) - U^0(t)\|_{H^s},$$

et par suite, puisque la série de terme général $L^n T^n / n!$ est convergente, que $(U^k)_{k \in \mathbb{N}}$ est de Cauchy dans $\mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$. Cet espace étant complet, on en déduit que $(U^k)_{k \in \mathbb{N}}$ converge, et sa limite U vérifie la formule intégrale (II.4) (en passant à la limite dans la relation de récurrence entre U^k et U^{k+1}). Supposons que U et V soient deux solutions fortes appartenant à \mathcal{E} . Alors en faisant la différence entre les deux formules intégrales,

$$\|U(t) - V(t)\|_{H^s} \leq L \int_0^t \|U(\tau) - V(\tau)\|_{H^s} d\tau,$$

d'où

$$\|U(t) - V(t)\|_{H^s} \leq e^{Lt} \|U(0) - V(0)\|_{H^s} = 0.$$

Ceci montre l'unicité des solutions fortes dans \mathcal{E} . Pour avoir l'unicité dans $\mathcal{C}(0, T; H^s(\mathbb{R}^d; \mathbb{R}^n))$ on utilise un argument de connexité. Si V est une autre solution forte (pour la donnée initiale $U_0 \in \mathcal{V}$), l'ensemble

$$\{t \in [0, T]; \forall s \in [0, t], \|V(t) - U(t)\|_{H^s} = 0\}$$

est fermé par construction, non vide car il contient $t = 0$, et ouvert d'après ce qui précède (en remplaçant U_* et U_0 par $U(t_0)$, et $t = 0$ par $t = t_0$). Enfin, pour montrer la continuité de l'application $U_0 \mapsto U$, on observe que si $U \in \mathcal{E}$ et $V \in \mathcal{E}$ sont des solutions fortes pour des données initiales $U_0 \in \mathcal{V}$ et $V_0 \in \mathcal{V}$ respectivement, on a pour tout $t \in [0, T]$,

$$\|U(t) - V(t)\|_{H^s} \leq \|S_t(U_0 - V_0)\|_{H^s} + L \int_0^t \|U(\tau) - V(\tau)\|_{H^s} d\tau,$$

d'où

$$\|U(t) - V(t)\|_{H^s} \leq e^{Lt} \|U_0 - V_0\|_{H^s}.$$

Donc l'application $U_0 \mapsto U$ est (localement) Lipschitzienne. ■

Remarque II.4

Les solutions de (II.4) ne sont pas a priori dérivables par rapport à t , car pour pouvoir dériver (en restant dans H^s) dans la formule intégrale (II.4) il faut que U_0 soit dans $\mathcal{D}(A)$ et que la fonction $t \mapsto F \circ U(t)$ appartienne à $L^1(0, T; \mathcal{D}(A))$. C'est possible au prix de l'augmentation de l'indice s selon l'ordre de l'opérateur A . En effet, si A est un opérateur différentiel d'ordre m , son domaine dans $H^s(\mathbb{R}^d)$ est $H^{s+m}(\mathbb{R}^d)$. Ainsi, si U est une solution forte dans $H^{s+m}(\mathbb{R}^d)$ (pour une donnée initiale $U_0 \in H^{s+m}(\mathbb{R}^d)$) on

$a \ U \in \mathcal{C}(0, T; H^{s+m}(\mathbb{R}^d)) \cap \mathcal{C}^1(0, T; H^s(\mathbb{R}^d))$ et pour tout $t \in [0, T]$,

$$U'(t) = AU(t) + F \circ U(t).$$

Quant aux équations de la forme (II.2), leur caractère bien posé (un non) n'est pas directement accessible par la méthode de type EDO utilisée pour (II.1), car la dérivation devant $F \circ U$ fait perdre un cran de régularité et empêche *a priori* de « boucler » le schéma itératif. Cependant, dans le cas du système de Boussinesq, on s'en sort en le reformulant comme

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \partial_x v \\ \partial_x u - \partial_{xxx}^3 u \end{pmatrix} + \begin{pmatrix} 0 \\ \partial_x(f(u) - u) \end{pmatrix}.$$

et en se plaçant dans l'espace $\mathcal{H} := H^s(\mathbb{R}) \times H^{s-1}(\mathbb{R})$, avec $s > 1/2$. La non-linéarité est localement Lipschitzienne dans cet espace puisque

$$\|\partial_x(f(u_1) - u_1) - \partial_x(f(u_2) - u_2)\|_{H^{s-1}} \lesssim \|f(u_1) - f(u_2)\|_{H^s} + \|u_1 - u_2\|_{H^s},$$

où le premier terme se majore grâce au théorème II.1. Et par ailleurs il s'avère que l'opérateur

$$A := \begin{pmatrix} 0 & \partial_x \\ \partial_x - \partial_{xxx}^3 & 0 \end{pmatrix},$$

de domaine $H^{s+2}(\mathbb{R}) \times H^{s+1}(\mathbb{R})$, est anti-autoadjoint pour le produit scalaire naturel sur \mathcal{H} , défini par

$$\left\langle \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} \right\rangle = \langle u_1, u_2 \rangle_{H^{s-1}} + \langle \partial_x u_1, \partial_x u_2 \rangle_{H^{s-1}} + \langle v_1, v_2 \rangle_{H^{s-1}}.$$

En effet, l'opérateur ∂_x étant anti-autoadjoint sur tous les H^s , on voit A est anti-symétrique : quels que soient $U_1 = (u_1, v_1)^t$, $U_2 = (u_2, v_2)^t \in \mathcal{D}(A) = H^{s+2}(\mathbb{R}) \times H^{s+1}(\mathbb{R})$,

$$\begin{aligned} \langle U_1, AU_2 \rangle &= \langle u_1, \partial_x v_2 \rangle + \langle \partial_x u_1, \partial_{xx}^2 v_2 \rangle + \langle v_1, \partial_x u_2 \rangle - \langle v_1, \partial_{xxx}^3 u_2 \rangle = \\ &= -\langle v_2, \partial_x u_1 \rangle + \langle v_2, \partial_{xxx}^3 u_1 \rangle - \langle u_2, \partial_x v_1 \rangle - \langle \partial_x u_2, \partial_{xx}^2 v_1 \rangle = -\langle U_2, AU_1 \rangle \end{aligned}$$

(où dans les égalités intermédiaires, \langle, \rangle désigne le produit scalaire sur H^{s-1}). On montre en outre facilement (comme A est un opérateur différentiel à coefficients constants) que A^* a pour domaine $\mathcal{D}(A) = H^{s+2}(\mathbb{R}) \times H^{s+1}(\mathbb{R})$. Par suite, d'après le théorème de Stone, A est le générateur infinitésimal d'un groupe unitaire (on peut d'ailleurs calculer son noyau de Green). Ainsi, en adaptant légèrement la démonstration du théorème II.2 (remplacer H^s par \mathcal{H} ; voir le théorème II.4) on montre que le problème de Cauchy associé au système de Boussinesq est localement bien posé dans \mathcal{H} , les solutions étant dans $\mathcal{C}(0, T; H^{s+2}(\mathbb{R}) \times H^{s+1}(\mathbb{R})) \cap \mathcal{C}^1(0, T; H^s(\mathbb{R}) \times H^{s-1}(\mathbb{R}))$ pour des données initiales dans $H^{s+2}(\mathbb{R}) \times H^{s+1}(\mathbb{R}) = \mathcal{D}(A)$.

Une théorie a été développée par Kato pour traiter des problèmes quasi-linéaires plus généraux.

Théorème II.3 (Kato)

On suppose que \mathcal{X} est un espace de Banach réflexif et que $\mathcal{Y} \hookrightarrow \mathcal{X}$ est dense dans \mathcal{X} . On suppose qu'il existe un isomorphisme S de \mathcal{Y} sur \mathcal{X} , et on équipe alors \mathcal{Y} de la

norme

$$\|y\|_{\mathcal{Y}} = \|Sy\|_{\mathcal{X}}.$$

On suppose ensuite que \mathcal{V} est une boule ouverte de \mathcal{Y} , et que pour tout $y \in \mathcal{V}$, $A(y)$ est un opérateur non borné sur \mathcal{X} , générateur d'un semi-groupe de contractions, tel que

- si y_0 est le centre de la boule \mathcal{V} , $A(y)y_0 \in \mathcal{Y}$ et est uniformément borné sur \mathcal{V} ,
- l'espace \mathcal{Y} est inclus dans le domaine de $A(y)$, $A(y)$ est continu de \mathcal{Y} dans \mathcal{X} , et l'application $y \mapsto A(y)$ est uniformément Lipschitzienne de \mathcal{V} , muni de la norme de \mathcal{X} , dans l'espace des opérateurs continus de \mathcal{Y} dans \mathcal{X} ,
- le domaine de $A(y)$ est tel que

$$x \in \mathcal{D}(A(y)) \iff (S^{-1}x \in \mathcal{D}(A(y)) \text{ et } A(y)S^{-1}x \in \mathcal{Y}),$$

l'opérateur $SA(y)S^{-1} - A$ est continu de \mathcal{X} dans \mathcal{X} , et sa norme est uniformément majorée sur \mathcal{V} .

On suppose enfin que $F : \mathcal{V} \rightarrow \mathcal{Y}$ est uniformément bornée, et uniformément \mathcal{X} -Lipschitzienne. Alors pour tout $U_0 \in \mathcal{V}$ le problème de Cauchy pour l'équation

$$U' = A(U)U + F \circ U, \quad U(0) = U_0$$

admet une unique solution forte locale en temps $U \in \mathcal{C}(0, T; \mathcal{Y}) \cap \mathcal{C}^1(0, T; \mathcal{X})$.

Ceci est un cas particulier de [9, théorème 6]. On peut en fait relaxer l'hypothèse sur les opérateurs $A(y)$, en demandant « seulement » qu'ils soient *m-quasi-accréatifs*, c'est-à-dire qu'ils soient générateurs de semi-groupes

$$(e^{tA(y)})_{t \geq 0}$$

tels que

$$\|e^{tA(y)}\| \leq e^{\beta t}$$

avec β uniforme sur \mathcal{V} (mais pas nécessairement $\beta \leq 0$).

Ceci permet notamment de montrer que l'équation de KdV est localement bien posée dans $H^s(\mathbb{R})$ avec $s \geq 3$. On pose en effet $\mathcal{X} = L^2(\mathbb{R})$, $\mathcal{Y} = H^s(\mathbb{R})$ avec $s \geq 3$, $S = \Lambda^s = (1 - \partial_{xx}^2)^{s/2}$ et $A(u) = -\partial_{xxx}^3 + f'(u)\partial_x$. Il n'est pas immédiat de montrer que $A(u)$ est m-quasi-accréatif. Cependant, cela se voit assez facilement pour $A_0(u) := f'(u)\partial_x$ (estimations *a priori* et résolution d'une équation de transport à coefficients variables, cf [1, Chap. 2]), et comme l'autre morceau $-\partial_{xxx}^3$ est générateur d'un groupe unitaire, on peut montrer (cf [9, pp. 501-502]) que cela implique que $A(u)$ est m-quasi-accréatif avec la même borne β que pour $A_0(u)$. Par ailleurs, on a

$$[S, A(u)]S^{-1} = [\Lambda^s, A_0(u)]\Lambda^{-s}$$

et il s'agit donc de montrer que $[\Lambda^s, A_0(u)]\Lambda^{-s}$ est un opérateur borné sur $L^2(\mathbb{R})$ lorsque $u \in H^s(\mathbb{R})$ avec $s \geq 3$. L'estimation du commutateur $[\Lambda^s, A_0(u)]$ n'est pas du tout triviale. Dans le cas $s = 2$ on pourrait faire le calcul explicite

$$[\Lambda^2, A_0(u)] = [\Lambda^2, f'(u)]\partial_x = -(f''(u)u_{xx} + f'''(u)u_x^2)\partial_x - 2f''(u)u_x\partial_{xx}^2$$

et l'on obtiendrait, en supposant quand même $u \in H^3(\mathbb{R})$,

$$\|[\Lambda^2, A_0(u)]\Lambda^{-2}v\|_{L^2} \leq \|f''(u)u_{xx} + f'''(u)u_x^2\|_{L^\infty}\|\partial_x\Lambda^{-2}v\|_{L^2} + 2\|f''(u)u_x\|_{L^\infty}\|\partial_{xx}^2\Lambda^{-2}v\|_{L^2}$$

$$\leq C(\|u\|_{H^3}) \|v\|_{L^2}$$

où C dépend continûment de son argument d'après l'injection de Sobolev $H^1(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$. Ce calcul se généralise (fastidieusement) aux nombres s qui sont des entiers pairs. Plus précisément, on peut montrer (voir par exemple [1, p. 470]) que pour $a, w \in H^s(\mathbb{R})$, $s > 1$, tels que $\partial_x a, \partial_x w \in L^\infty(\mathbb{R})$,

$$\|[\partial_x^n, a\partial_x]w\|_{L^2} \lesssim \|\partial_x a\|_{L^\infty} \|w\|_{H^s} + \|\partial_x w\|_{L^\infty} \|a\|_{H^s}$$

pour tout entier $n \leq s$. Ainsi, si s est un entier pair (supérieur à 3),

$$\begin{aligned} \|[\Lambda^s, f'(u)\partial_x]\Lambda^{-s}v\|_{L^2} &\lesssim \|\partial_x f'(u)\|_{L^\infty} \|v\|_{L^2} + \|\partial_x \Lambda^{-s}v\|_{L^\infty} \|f'(u)\|_{H^s} \\ &\lesssim \|f'(u)\|_{H^s} \|v\|_{L^2} \end{aligned}$$

grâce à l'injection de Sobolev $H^{s-1}(\mathbb{R}) \hookrightarrow L^\infty(\mathbb{R})$. Le terme $\|f'(u)\|_{H^s}$ se majore ensuite en fonction de $\|u\|_{H^s}$ grâce au théorème II.1 (on peut toujours se ramener à $f'(0) = 0$ en remarquant que $[\Lambda^s, f'(u)\partial_x] = [\Lambda^s, (f'(u) - f'(0)\partial_x)]$).

2.3 Solutions moins régulières

Lorsque l'on considère des données initiales moins régulières, typiquement dans $H^s(\mathbb{R}^d)$ avec $s \leq d/2$, le schéma de preuve utilisé pour le théorème II.2 ne fonctionne plus. Cependant, on peut espérer l'adapter si la perte de régularité due à la non-linéarité F est « récupérée » par l'opérateur de Duhamel

$$(II.5) \quad D : f \mapsto Df; \quad Df(t) = \int_0^t S_{t-\tau} f(\tau) d\tau.$$

D'un point de vue abstrait cela peut s'énoncer comme suit, où \mathcal{S} joue le rôle d'un espace de fonctions plus régulières que celles de l'espace \mathcal{N} .

Théorème II.4

Soient \mathcal{S} et \mathcal{N} deux espaces de Banach. On suppose que l'on a

1. un opérateur linéaire continu $D : \mathcal{N} \rightarrow \mathcal{S}$ de norme C_0 ,
2. un opérateur non-linéaire $N : \mathcal{S} \rightarrow \mathcal{N}$, tel que $N(0) = 0$, et il existe $\varepsilon > 0$ tel que N soit Lipschitzienne de rapport $1/(2C_0)$ dans la boule

$$B_\varepsilon := \{u \in \mathcal{S}; \|u\| \leq \varepsilon\}.$$

Alors, pour tout $u_\ell \in B_{\varepsilon/2}$, il existe un unique $u \in B_\varepsilon$ solution de l'équation implicite

$$u = u_\ell + DN(u),$$

En outre, l'application $u \in B_\varepsilon \mapsto u$ est Lipschitzienne de rapport 2.

Démonstration : Sans originalité, on construit une solution par un schéma itératif

$$u^0 = u_\ell, \quad u^{n+1} = u_\ell + DN(u^n).$$

Les hypothèses sont faites pour que $u^n \in B_\varepsilon$ si $u_\ell \in B_{\varepsilon/2}$, et pour que

$$\|u^{n+1} - u^n\|_{\mathcal{S}} \leq \frac{1}{2} \|u^n - u^{n-1}\|_{\mathcal{S}}.$$

Donc la suite (u^n) est convergente dans \mathcal{S} . Sa limite u appartient à B_ε et vérifie l'équation $u = u_\ell + DN(u)$. Si jamais v était une autre solution dans B_ε on aurait

$$\|u - v\|_{\mathcal{S}} \leq \frac{1}{2} \|u - v\|_{\mathcal{S}}$$

et donc $u = v$. Enfin on obtient par différence de $u = u_\ell + DN(u)$ et $v = v_\ell + DN(v)$ que

$$\|u - v\|_{\mathcal{S}} \leq 2 \|u_\ell - v_\ell\|_{\mathcal{S}}.$$

■

On peut espérer appliquer ce résultat abstrait à un opérateur de Duhamel D défini par (II.5) à l'aide d'un groupe unitaire (S_t) sur un espace de Hilbert \mathcal{H} , si l'on dispose d'un espace \mathcal{S} s'injectant continûment dans $\mathcal{C}(0, T; \mathcal{H})$ et d'un espace \mathcal{N} tel que $D : \mathcal{N} \rightarrow \mathcal{S}$.

Corollaire II.1

Si A est un opérateur anti-autoadjoint sur un espace de Hilbert \mathcal{H} , si \mathcal{S}_T et \mathcal{N}_T sont deux espaces de Banach tels que \mathcal{S}_T s'injecte continûment dans $\mathcal{C}(0, T; \mathcal{H})$, si l'opérateur de Duhamel (II.5) associé au groupe engendré par A est continu de \mathcal{N}_T dans \mathcal{S}_T , si $N : \mathcal{S}_T \rightarrow \mathcal{N}_T$, tel que $N(0) = 0$, est Lipschitzienne de rapport $1/(2\|D\|)$ (pour T assez petit) dans la boule

$$B_\varepsilon := \{u \in \mathcal{S}_T; \|u\| \leq \varepsilon\},$$

alors le problème de Cauchy pour

$$\partial_t U = AU + N(U)$$

est localement conditionnellement bien posé dans \mathcal{H} (la condition d'unicité étant portée par l'ensemble $\mathcal{E} = \mathcal{S}_T$) au voisinage de 0.

Dans les « bons » cas, ce sont des estimations de Strichartz qui permettent d'obtenir la « régularisation » par l'opérateur de Duhamel. On peut ainsi énoncer :

Meta-théorème II.1

Soit une EPD de la forme (II.1) avec A un opérateur anti-autoadjoint sur $L^2(\mathbb{R}^d; \mathbb{R}^n)$ et $N : U \mapsto F \circ U$ avec $F \in \mathcal{C}^\infty(\mathbb{R}^n; \mathbb{R}^n)$ une non-linéarité pour laquelle $L^2(\mathbb{R}^d; \mathbb{R}^n)$ est sous-critique. On suppose que le semi-groupe engendré par A , et l'opérateur de Duhamel associé (II.5), vérifient des estimations de dispersion et de Strichartz pour un ensemble compact de couples (q, r) , et l'on définit l'espace $\mathcal{S}^0([0, T] \times \mathbb{R}^d; \mathbb{R}^n) \subset \mathcal{C}([0, T]; L^2(\mathbb{R}^d; \mathbb{R}^n))$ comme l'adhérence de $\mathcal{S}([0, T] \times \mathbb{R}^d; \mathbb{R}^n)$ pour la norme

$$\max_{(q,r)} \|U\|_{L^q(0,T;L^r(\mathbb{R}^d;\mathbb{R}^n))}.$$

Alors le problème de Cauchy est localement conditionnellement bien posé dans $L^2(\mathbb{R}^d; \mathbb{R}^n)$, la condition d'unicité étant portée par l'ensemble $\mathcal{E} = \mathcal{S}^0([0, T] \times \mathbb{R}^d; \mathbb{R}^n)$.

Pour en faire un vrai résultat nous allons considérer le cas de (NLS) avec une non-linéarité algébrique $f(u) = \mu|u|^{p-1}u$, où p un entier impair (ce qui assure la régularité de f). On constate qu'il y a très peu de choix d'entiers p possibles pour que $L^2(\mathbb{R}^d)$ soit sous-critique. En effet,

$$0 > d/2 - 2/(p-1) \quad \Leftrightarrow \quad 1 < p < 1 + 4/d.$$

Dans le cas $d = 1$ on trouve $p = 3$, et dès la dimension $d = 2$, $L^2(\mathbb{R}^d)$ est critique pour $p = 3$, sur-critique pour $p \geq 5$. (En dimension $d = 3$, L^2 est sur-critique pour $p = 3$; il faut aller chercher H^1 pour avoir un espace sous-critique.) Ainsi la portée du « meta-théorème » II.1 est assez limitée. Concernant (NLS) il conduit au seul résultat suivant.

Théorème II.5

Considérons l'équation

$$(II.6) \quad i\partial_t u + \partial_{xx}^2 u = \mu |u|^2 u$$

avec $\mu = \pm 1$. Pour tout $T > 0$ on définit $\mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C}) \subset \mathcal{C}([0, T]; L^2(\mathbb{R}; \mathbb{C}))$ comme l'adhérence de $\mathcal{S}([0, T] \times \mathbb{R}; \mathbb{C})$ pour la norme

$$\|u\|_{\mathcal{S}^0} := \max_{(q,r) \in \Sigma} \|u\|_{L^q(0,T; L^r(\mathbb{R}^d; \mathbb{C}))}, \quad \Sigma := \{(q, r); 2/q + 1/r = 1/2, 2 \leq q, r \leq +\infty\}.$$

Pour tout $R > 0$ on note

$$B_R := \{v \in L^2(\mathbb{R}; \mathbb{C}); \|v\|_{L^2(\mathbb{R}; \mathbb{C})} \leq R\}.$$

Alors pour tout $R > 0$ il existe $T > 0$ tel que pour tout $u_0 \in B_R$ il existe un unique $u \in \mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ qui soit solution forte dans $L^2(\mathbb{R}; \mathbb{C})$ de (II.6) pour la donnée initiale u_0 . De plus, l'application $u_0 \in B_R \mapsto u \in \mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ est Lipschitzienne.

Démonstration : On rappelle (voir théorème I.4 et les précisions qui le suivent) que pour $(q, r) \in \Sigma$, la solution $u(t) = S_t u_0$ de l'équation de Schrödinger libre

$$i\partial_t u + \partial_{xx}^2 u = 0$$

vérifie l'estimation

$$\|u\|_{L^q(I; L^r(\mathbb{R}; \mathbb{C}))} \lesssim \|u(0)\|_{L^2(\mathbb{R}; \mathbb{C})}.$$

On a également une « estimation duale » (dans l'argument T^*T elle provient de la norme de T , tandis que l'estimation précédente provenait de la norme de T^*) :

$$\left\| \int_I S_{-s} f(s) \, ds \right\|_{L^2(\mathbb{R}; \mathbb{C})} \lesssim \|f\|_{L^{q'}(I; L^{r'}(\mathbb{R}; \mathbb{C}))},$$

et

$$\left\| t \mapsto \int_I S_{t-s} f(s) \, ds \right\|_{L^q(I; L^r(\mathbb{R}; \mathbb{C}))} \lesssim \|f\|_{L^{q'}(I; L^{r'}(\mathbb{R}; \mathbb{C}))},$$

qui provient de la norme de T^*T . Dans toutes ces estimations I est un intervalle quelconque d'intégration en temps, par exemple \mathbb{R}^+ . Mais en fait on a surtout besoin d'une estimation de Strichartz dite « retardée », c'est-à-dire de

$$\left\| t \mapsto \int_0^t S_{t-s} f(s) \, ds \right\|_{L^q(\mathbb{R}; L^r(\mathbb{R}; \mathbb{C}))},$$

ce qui revient à estimer les solutions de l'équation avec terme source

$$(II.7) \quad i\partial_t u + \partial_{xx}^2 u = f$$

(qui s'annulent en $t = 0$). Une telle estimation peut se déduire de celle de $\int_0^{+\infty} S_{t-s} f(s) \, ds$ grâce au lemme II.1 énoncé ci-après. Il s'applique à $K(t, s) = S_{t-s}$, qui est continu de

$\mathcal{X} = L^{r'}$ dans $\mathcal{Y} = L^r$ pour $r \geq 2$ d'après (I.3), et $p = q'$: noter qu'on a bien $p < q$ puisqu'en fait $q \in [4, +\infty]$ pour $(q, r) \in \Sigma$. Par suite les solutions de (II.7) vérifient l'estimation a priori, pour tout couple $(q, r) \in \Sigma$,

$$\|u\|_{L^q(0,T;L^r(\mathbb{R};\mathbb{C}))} \leq C(T) \left(\|u(0)\|_{L^2(\mathbb{R};\mathbb{C})} + \|f\|_{L^{q'}(0,T;L^{r'}(\mathbb{R};\mathbb{C}))} \right),$$

où C dépend continûment de T (et ne dépend pas de (q, r)). En particulier, l'opérateur de Duhamel (II.5) vérifie l'estimation a priori

$$\|Df\|_{L^q(0,T;L^r(\mathbb{R};\mathbb{C}))} \leq C(T) \|f\|_{L^{q'}(0,T;L^{r'}(\mathbb{R};\mathbb{C}))}$$

pour tout $(q, r) \in \Sigma$. En fait, on a une estimation beaucoup plus générale, à savoir

$$\|Df\|_{L^{\tilde{q}}(0,T;L^{\tilde{r}}(\mathbb{R};\mathbb{C}))} \leq C(T) \|f\|_{L^{q'}(0,T;L^{r'}(\mathbb{R};\mathbb{C}))}$$

pour tous couples (q, r) et $(\tilde{q}, \tilde{r}) \in \Sigma$. Ceci découle du lemme II.2 montré ci-après. Reste à savoir si, au moins pour un certain couple $(q, r) \in \Sigma$ et quitte à diminuer T , la non-linéarité $u \mapsto |u|^2 u$ est Lipschitzienne de $L^q(0, T; L^r(\mathbb{R}; \mathbb{C}))$ dans $L^{q'}(0, T; L^{r'}(\mathbb{R}; \mathbb{C}))$ avec un rapport assez petit. Soit $r = 4$, de sorte que $r' = 4/3 = r/3$, $q = 8$, $q' = 8/7$. Par l'inégalité de Hölder en temps

$$\||u|^2 u - |v|^2 v\|_{L^{q'}(0,T;L^{r'}(\mathbb{R};\mathbb{C}))} \leq T^{1/2} \||u|^2 u - |v|^2 v\|_{L^{q/3}(0,T;L^{r/3}(\mathbb{R};\mathbb{C}))},$$

et en majorant de façon élémentaire (par la formule de Taylor avec reste intégral)

$$||u|^2 u - |v|^2 v| \leq 2 |u - v| (|u|^2 + |v|^2),$$

on obtient par l'inégalité Hölder en espace-temps,

$$\||u|^2 u - |v|^2 v\|_{L^{q'}(0,T;L^{r'}(\mathbb{R};\mathbb{C}))} \leq$$

$$2 T^{1/2} \|u - v\|_{L^q(0,T;L^r(\mathbb{R};\mathbb{C}))} (\|u\|_{L^q(0,T;L^r(\mathbb{R};\mathbb{C}))}^2 + \|v\|_{L^q(0,T;L^r(\mathbb{R};\mathbb{C}))}^2).$$

■

Lemme II.1 (*Christ-Kiselev, 2001*)

Soient \mathcal{X}, \mathcal{Y} deux espaces de Banach, I un intervalle de \mathbb{R} , et $K \in \mathcal{C}(I \times I; \mathcal{B}(\mathcal{X}; \mathcal{Y}))$ (où $\mathcal{B}(\mathcal{X}; \mathcal{Y})$ désigne l'espace des opérateurs linéaires continus de \mathcal{X} dans \mathcal{Y}). On suppose que pour $1 \leq p < q \leq +\infty$,

$$\begin{aligned} \mathcal{K} : L^p(I; \mathcal{X}) &\rightarrow L^q(I; \mathcal{Y}) \\ f &\mapsto \mathcal{K}f; \quad \mathcal{K}f(t) = \int_I K(t, s) f(s) \, ds \end{aligned}$$

est continu de norme C . Alors on a, pour tout $f \in L^p(I; \mathcal{X})$,

$$\left\| t \mapsto \int_{s \in I; s < t} K(t, s) f(s) \, ds \right\|_{L^q(I; \mathcal{Y})} \leq C \|f\|_{L^p(I; \mathcal{X})}.$$

Avant la publication de ce lemme les estimations de Strichartz retardées étaient traitées au cas par cas (cf [7], §3).

Lemme II.2 (Keel-Tao, 1997)

Soit une famille à un paramètre $(S_t)_{t \in \mathbb{R}}$ d'opérateurs unitaires sur $L^2(\mathbb{R})$ vérifiant les estimations a priori suivantes

$$(II.8) \quad \left\| \int_I S_s^* f(s) \, ds \right\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^{q'}(I; L^{r'}(\mathbb{R}))}$$

pour tout intervalle I , et

$$(II.9) \quad \left\| t \mapsto \int_0^t S_t S_s^* f(s) \, ds \right\|_{L^q(0, T; L^r(\mathbb{R}))} \lesssim \|f\|_{L^{q'}(0, T; L^{r'}(\mathbb{R}))},$$

pour tout couple $(q, r) \in \Sigma$, c'est-à-dire tel que $(1/q, 1/r)$ appartienne au segment d'extrémités $(0, 1/2)$ et $(1/4, 0)$. Alors cette famille vérifie l'estimation

$$(II.10) \quad \left\| t \mapsto \int_0^t S_t S_s^* f(s) \, ds \right\|_{L^{\tilde{q}}(0, T; L^{\tilde{r}}(\mathbb{R}))} \lesssim \|f\|_{L^{q'}(0, T; L^{r'}(\mathbb{R}))},$$

quels que soient les couples (q, r) et (\tilde{q}, \tilde{r}) dans Σ .

Démonstration : (Esquisse de...) Par dualité, (II.10) équivaut à

$$\left| \int_0^T \int_0^t \langle S_t^* g(t), S_s^* f(s) \rangle \, ds \, dt \right| \lesssim \|f\|_{L^{q'}(0, T; L^{r'}(\mathbb{R}))} \|g\|_{L^{\tilde{q}'}(0, T; L^{\tilde{r}'}(\mathbb{R}))}.$$

Or d'après (II.9) cette estimation est vraie pour $(q, r) = (\tilde{q}, \tilde{r})$ elle l'est aussi pour $(q, r) = (\infty, 2)$ ou $(\tilde{q}, \tilde{r}) = (\infty, 2)$. En effet,

$$\begin{aligned} \left| \int_0^T \int_0^t \langle S_t^* g(t), S_s^* f(s) \rangle \, ds \, dt \right| &= \left| \int_0^T \langle S_t^* g(t), \int_0^t S_s^* f(s) \, ds \rangle \, dt \right| \\ &\leq \|S^* g\|_{L^1(0, T; L^2(\mathbb{R}))} \sup_{t \in [0, T]} \left\| \int_0^t S_s^* f(s) \, ds \right\|_{L^2(\mathbb{R})} \lesssim \|g\|_{L^1(0, T; L^2(\mathbb{R}))} \|f\|_{L^{q'}(0, T; L^{r'}(\mathbb{R}))} \end{aligned}$$

d'après (II.8) et le fait que S_t est unitaire sur $L^2(\mathbb{R})$. De façon analogue,

$$\begin{aligned} \left| \int_0^T \int_0^t \langle S_t^* g(t), S_s^* f(s) \rangle \, ds \, dt \right| &= \left| \int_0^T \langle \int_s^T S_t^* g(t) \, dt, S_s^* f(s) \rangle \, ds \right| \\ &\leq \|S^* f\|_{L^1(0, T; L^2(\mathbb{R}))} \sup_{s \in [0, T]} \left\| \int_s^T S_t^* g(t) \, dt \right\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^1(0, T; L^2(\mathbb{R}))} \|g\|_{L^{\tilde{q}'}(0, T; L^{\tilde{r}'}(\mathbb{R}))}. \end{aligned}$$

On peut ensuite conclure grâce à un résultat d'*interpolation bilinéaire* (cf [2, p. 76], Exercice 3.13.5) et aux résultats classiques sur l'interpolation des espaces L^q (cf [2, p. 109]) et $L^q(0, T; L^r(\mathbb{R}))$ (plus généralement pour les espaces L^q à valeurs dans des espaces de Banach, cf [2, p. 130]). ■

Le théorème II.5 ne fait qu'effleurer le champ extrêmement vaste de l'état de l'art concernant (NLS). Pour plus de résultats on renvoie par exemple à [14, ch. 3]. Nous allons dans la fin de ce chapitre aborder la question de l'existence de solutions *globales* en temps, où l'on verra le rôle que peut jouer le caractère *défocalisant* de la non-linéarité.

Avant cela, il faut signaler que les estimations de Strichartz sont surtout utiles pour traiter les équations semi-linéaires. Lorsque la non-linéarité implique une dérivée, comme dans l'équation de KdV, le problème est plus délicat et nécessite d'autres estimations a priori, qui permettent vraiment de gagner en régularité et pas seulement en intégrabilité.

Lemme II.3

Les solutions (suffisamment régulières) de l'équation d'Airy avec terme source

$$(II.11) \quad \partial_t u + \partial_{xxx}^3 u = f$$

vérifient l'estimation a priori

$$(II.12) \quad \|\partial_x u\|_{L^\infty(\mathbb{R}; L^2(0, T))} \lesssim \|u(0)\|_{L^2(\mathbb{R})} + \|f\|_{L^1(0, T; L^2(\mathbb{R}))},$$

et plus généralement, pour tout $s \in \mathbb{R}$,

$$(II.13) \quad \|\Lambda^s \partial_x u\|_{L^\infty(\mathbb{R}; L^2(0, T))} \lesssim \|u(0)\|_{H^s(\mathbb{R})} + \|f\|_{L^1(0, T; H^s(\mathbb{R}))}.$$

Démonstration : On remarque qu'il suffit de montrer l'estimation

$$\|\partial_x u(\cdot, 0)\|_{L^2(0, T)} \lesssim \|u(0)\|_{L^2(\mathbb{R})} + \|f\|_{L^1(0, T; L^2(\mathbb{R}))},$$

car pour tout $x_0 \in \mathbb{R}$, l'application $\tilde{u} : (t, x) \mapsto u(t, x - x_0)$ est solution de l'équation $\partial_t \tilde{u} + \partial_{xxx}^3 \tilde{u} = \tilde{f}$ où $\tilde{f} : (t, x) \mapsto f(t, x - x_0)$, et

$$\|u(0)\|_{L^2(\mathbb{R})} = \|\tilde{u}(0)\|_{L^2(\mathbb{R})}, \quad \|f\|_{L^1(0, T; L^2(\mathbb{R}))} = \|\tilde{f}\|_{L^1(0, T; L^2(\mathbb{R}))}.$$

On commence par traiter le cas $f = 0$, et l'on montre que

$$(II.14) \quad \int_{-\infty}^{+\infty} (\partial_x u(t, 0))^2 dt \lesssim \|u(0)\|_{L^2(\mathbb{R})}^2.$$

En effet, on sait que la transformée de Fourier de u en espace est donnée par

$$\widehat{u}(t, 0) = \widehat{u}(0, \xi) e^{i\xi^3 t}.$$

D'autre part, soit \check{u} la transformée de Fourier de u en temps. Par le théorème de Plancherel on a

$$\int_{-\infty}^{+\infty} (\partial_x u(t, 0))^2 dt = \int_{-\infty}^{+\infty} (\partial_x \check{u}(\tau, 0))^2 d\tau,$$

et pour tout $\phi \in \mathcal{S}(\mathbb{R}; \mathbb{R})$,

$$\int_{-\infty}^{+\infty} \phi(\tau) \partial_x \check{u}(\tau, 0) d\tau = \int_{-\infty}^{+\infty} \check{\phi}(t) \partial_x u(t, 0) dt = \iint \check{\phi}(t) i \xi \widehat{u}(0, \xi) e^{i\xi^3 t} dt d\xi.$$

Par les théorèmes de Fubini, Plancherel et Cauchy-Schwarz, on a donc

$$\begin{aligned} \left| \int_{-\infty}^{+\infty} \phi(\tau) \partial_x \check{u}(\tau, 0) d\tau \right| &= \left| \int i \xi \widehat{u}(0, \xi) \int \check{\phi}(t) e^{i\xi^3 t} dt d\xi \right| = \left| \int 2\pi i \xi \widehat{u}(0, \xi) \phi(\xi^3) d\xi \right| \\ &\leq 2\pi \|u(0, \cdot)\|_{L^2(\mathbb{R})} \left(\int \xi^2 \phi(\xi^3)^2 d\xi \right)^{1/2} = \pi \|u(0, \cdot)\|_{L^2(\mathbb{R})} \|\phi\|_{L^2(\mathbb{R})}. \end{aligned}$$

Ceci prouve (II.14). On en déduit (II.12) grâce à la formule de Duhamel. En effet, en posant $z(t, x, s) = \partial_x(S_{t-s}f(s))$, on a pour tout $s \in \mathbb{R}$ et pour presque tout $x \in \mathbb{R}$,

$$\|z(\cdot, x, s)\|_{L^2(\mathbb{R})} \lesssim \|f(s)\|_{L^2(\mathbb{R})}.$$

Soit $Z(t, x) = \int_0^t z(t, x, s) ds$. Pour $t \in [0, T]$ on a

$$|Z(t, x)| \leq \int_0^T |z(t, x, s)| ds$$

d'où

$$\|Z(\cdot, x)\|_{L^2(0, T)} \leq \int_0^T \|z(\cdot, x, s)\|_{L^2(\mathbb{R})} ds \lesssim \|f\|_{L^1(0, T; L^2(\mathbb{R}))}. \quad \blacksquare$$

Si l'on considère l'équation de KdV ordinaire, avec un non-linéarité quadratique, le résultat d'existence locale de solutions fortes dans H^s avec $s \geq 3$ montré à l'aide du théorème abstrait de Kato est loin du cas critique $s = -3/2$. Le lemme II.3 et le résultat suivant (que l'on admettra) permettent de descendre plus près du cas critique, jusqu'à $s > 3/4$.

Lemme II.4 (*Kenig-Ponce-Vega*)

Les solutions (suffisamment régulières) de l'équation d'Airy avec terme source (II.11) vérifient l'estimation a priori

$$(II.15) \quad \|u\|_{L^2(\mathbb{R}; L^\infty(0, T))} \lesssim \|u(0)\|_{H^s(\mathbb{R})} + \|f\|_{L^1(0, T; H^s(\mathbb{R}))},$$

pour tout $s > 3/4$.

L'intérêt de cette estimation se voit lorsqu'on substitue $u\partial_x u$ à f dans (II.13) : pour pouvoir « boucler » l'estimation a priori pour l'équation (non-linéaire) de KdV, il faut contrôler $\|u\partial_x u\|_{L^1(0, T; H^s(\mathbb{R}))}$ et donc en particulier $\|u\Lambda^s \partial_x u\|_{L^1(0, T; L^2(\mathbb{R}))}$. Or par l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \|u\Lambda^s \partial_x u\|_{L^1(0, T; L^2(\mathbb{R}))} &\leq T^{1/2} \|u\Lambda^s \partial_x u\|_{L^2(0, T; L^2(\mathbb{R}))} = T^{1/2} \|u\Lambda^s \partial_x u\|_{L^2(\mathbb{R}; L^2(0, T))} \\ &\leq T^{1/2} \|\Lambda^s \partial_x u\|_{L^\infty(\mathbb{R}; L^2(0, T))} \|u\|_{L^2(\mathbb{R}; L^\infty(0, T))}. \end{aligned}$$

Théorème II.6 (*Kenig-Ponce-Vega*)

L'équation de KdV avec une non-linéarité quadratique est localement bien posée dans $H^s(\mathbb{R})$ pour tout $s > 3/4$.

2.4 Solutions globales

Les résultats aperçus précédemment donnent des solutions (fortes) locales en temps dans H^s (pour certaines valeurs de s). Une question naturelle est ensuite de savoir dans quels cas ces solutions peuvent être globales. Or, (comme pour les équations différentielles ordinaires) pour un problème localement bien posé dans H^s , si une solution maximale n'est pas prolongeable au delà d'un temps fini $T > 0$, c'est que $\|u(t)\|_{H^s}$ tend vers l'infini lorsque $t \nearrow T$. Montrer que l'on a des solutions globales demande donc de « contrôler » la norme de ces solutions dans $\mathcal{C}(0, t; H^s)$ (qui pourrait « exploser » lorsque $t \nearrow T$). Pour NSL avec une non-linéarité algébrique par exemple, il s'avère que la norme $L^{p-1}(0, t; L^\infty(\mathbb{R}^d))$ permet ce contrôle.

Proposition II.1

Soient $s \geq 0$, p un entier impair. Il existe $C > 0$ tel que si $t > 0$ et $u \in \mathcal{C}(0, t; H^s(\mathbb{R}^d))$ est une solution forte dans $H^s(\mathbb{R}^d)$

$$(II.16) \quad i\partial_t u + \partial_{xx}^2 u = \mu |u|^{p-1} u,$$

alors

$$\|u\|_{\mathcal{C}(0,t;H^s(\mathbb{R}^d))} \leq \|u(0)\|_{H^s(\mathbb{R}^d)} \exp(C\|u\|_{L^{p-1}(0,t;L^\infty(\mathbb{R}^d))}^p).$$

Démonstration : D'après la formule de Duhamel,

$$\begin{aligned} \|u(t)\|_{H^s(\mathbb{R}^d)} &\leq \|u(0)\|_{H^s(\mathbb{R}^d)} + |\mu| \int_0^t \| |u(\tau)|^{p-1} u(\tau) \|_{H^s(\mathbb{R}^d)} d\tau \\ &\leq \|u(0)\|_{H^s(\mathbb{R}^d)} + |\mu| \int_0^t \|u(\tau)\|_{L^\infty(\mathbb{R}^d)}^{p-1} \|u(\tau)\|_{H^s(\mathbb{R}^d)} d\tau \end{aligned}$$

d'après le théorème II.1. On conclut par le lemme de Gronwall. ■

Ce résultat montre la *persistance de régularité* des solutions : les temps d'existence de solutions fortes dans des espaces H^s d'indice différents sont les mêmes ! Il permet notamment de justifier de lois de conservation ; par exemple, la conservation de la norme L^2 pour les solutions fortes de l'équation (II.6).

Proposition II.2

Soit $u \in \mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ une solution forte dans $L^2(\mathbb{R})$ de (II.6). Alors $\|u(t)\|_{L^2(\mathbb{R})}$ est indépendante de t .

Démonstration : On peut invoquer un argument de connexité pour se ramener à un problème local. En effet, l'ensemble

$$A := \{t \in [0, T]; \forall s \leq t, \|u(t)\|_{L^2(\mathbb{R})} = \|u(0)\|_{L^2(\mathbb{R})}\}$$

contient 0 et il est fermé puisque $u \in \mathcal{C}(0, T; L^2(\mathbb{R}))$. Pour montrer qu'il est égal à $[0, T]$ il reste à montrer qu'il est ouvert dans $[0, T]$. Il suffit par translation de montrer qu'il existe $t > 0$ tel que $[0, t] \subset A$. Si on savait que $u \in \mathcal{S}([0, T] \times \mathbb{R}; \mathbb{C})$, on pourrait faire le calcul suivant :

$$\begin{aligned} \frac{d}{dt} \int |u(t, \mathbf{x})|^2 d\mathbf{x} &= 2\operatorname{Re} \int \overline{u(t, \mathbf{x})} \partial_t u(t, \mathbf{x}) d\mathbf{x} = \\ 2\operatorname{Re} \left(i \int \overline{u(t, \mathbf{x})} \Delta u(t, \mathbf{x}) + \mu |u(t, \mathbf{x})|^3 d\mathbf{x} \right) &= 0. \end{aligned}$$

Pour u « seulement » dans $\mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ ce calcul n'est pas justifié. Cependant, on peut approcher u_0 dans $L^2(\mathbb{R})$ par une suite d'éléments u_0^n de $H^s(\mathbb{R})$ avec s assez grand. D'après le théorème II.5, il existe $t > 0$ et pour tout $n \in \mathbb{N}$, une solution forte $u^n \in \mathcal{S}^0([0, t] \times \mathbb{R}; \mathbb{C})$ de (II.6) pour la donnée initiale u_0^n . De plus, u^n converge vers u dans $\mathcal{S}^0([0, t] \times \mathbb{R}; \mathbb{C})$, et en particulier dans $\mathcal{C}(0, t; L^2(\mathbb{R}))$. Pour montrer que $[0, t] \subset A$, il suffit donc de montrer que $\|u^n(\tau)\|_{L^2(\mathbb{R})}$ est indépendante de τ . Or d'après la proposition II.1,

$$\|u^n\|_{\mathcal{C}(0,t;H^s(\mathbb{R}^d))} \leq \|u(0)\|_{H^s(\mathbb{R}^d)} \exp(C\|u^n\|_{L^2(0,t;L^\infty(\mathbb{R}^d))}^3).$$

Et puisque u^n converge dans $\mathcal{S}^0([0, t] \times \mathbb{R}; \mathbb{C})$, $\|u^n\|_{L^4(0, t; L^\infty(\mathbb{R}))}$ est bornée, et donc $\|u^n\|_{L^2(0, t; L^\infty(\mathbb{R}))}$ aussi. Donc u^n appartient à $\mathcal{C}(0, t; H^s(\mathbb{R}^d))$ et donc, sans pour autant avoir $u^n \in \mathcal{S}([0, T] \times \mathbb{R}; \mathbb{C})$, on peut avoir $u^n \in \mathcal{C}^2([0, T] \times \mathbb{R}; \mathbb{C})$ en prenant s assez grand au départ. ■

Théorème II.7

Soit $u_0 \in L^2(\mathbb{R})$. Quel que soit $T > 0$ il existe un unique $u \in \mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ solution forte dans $L^2(\mathbb{R}; \mathbb{C})$ de (II.6) pour la donnée initiale u_0 . En outre l'application $u_0 \mapsto u \in \mathcal{S}^0([0, T] \times \mathbb{R}; \mathbb{C})$ est continue. Autrement dit, (II.6) est globalement bien posée dans $L^2(\mathbb{R})$.

Démonstration : Appliquer le théorème II.5 de façon répétée sur une subdivision de $[0, T]$ en intervalles assez petits en fonction de $\|u_0\|_{L^2(\mathbb{R})}$. ■

On peut parfois utiliser une autre loi de conservation, comme celle de l'énergie dans les cas défocalisants ($\mu = 1$),

$$E[u] := \int \frac{1}{2} |\nabla u(t, \mathbf{x})|^2 + \frac{2}{p+1} |u(t, \mathbf{x})|^{p+1} d\mathbf{x},$$

qui contrôle évidemment la norme H^1 .

Théorème II.8

L'équation NLS cubique défocalisante en dimension 2 est globalement bien posée dans $H^1(\mathbb{R}^2)$.

Citons pour finir un résultat global pour KdV.

Théorème II.9 (Kenig-Ponce-Vega)

L'équation de KdV avec une non-linéarité quadratique est globalement bien posée dans $H^s(\mathbb{R})$ quel que soit $s \geq 1$.

Chapitre III

Symétries et lois de conservation

Dans ce chapitre on va s'intéresser plus particulièrement aux propriétés algébriques des EDP rencontrées précédemment.

1 Groupes de symétrie des EDP

Commençons par quelques définitions très générales, dans lesquelles $\Omega \times U$ désigne un ouvert de $\mathbb{R}^d \times \mathbb{R}^n$.

Définition III.1

Un groupe local de transformations sur $\Omega \times U$ est donné par l'action d'un groupe de Lie G (c'est-à-dire un groupe par exemple multiplicatif qui soit aussi une variété différentiable, de telle sorte que les applications $(g, h) \mapsto gh$ et $g \mapsto g^{-1}$ soient régulières), qui à tout g voisin de l'élément neutre e de G , et à tout $(\mathbf{x}, \mathbf{u}) \in \Omega \times U$ associe un élément $g \cdot (\mathbf{x}, \mathbf{u})$ en respectant les propriétés suivantes :

- quel que soit $(\mathbf{x}, \mathbf{u}) \in \Omega \times U$, $e \cdot (\mathbf{x}, \mathbf{u}) = (\mathbf{x}, \mathbf{u})$,
- quel que soit $(\mathbf{x}, \mathbf{u}) \in \Omega \times U$, pour tous g, h voisins de e dans G , $g \cdot (h \cdot (\mathbf{x}, \mathbf{u})) = (gh) \cdot (\mathbf{x}, \mathbf{u})$.

Lorsque G est de dimension 1, on dit qu'on a un groupe de transformations à un paramètre. On définit son générateur infinitésimal comme le champ de vecteurs sur $\Omega \times U$:

$$\mathbf{v}(\mathbf{x}, \mathbf{u}) = \frac{d}{dg} (g \cdot (\mathbf{x}, \mathbf{u}))|_{g=e} \in \mathbb{R}^d \times \mathbb{R}^n.$$

Il est d'usage de noter

$$\mathbf{v} = \sum_{i=1}^d \xi_i \frac{\partial}{\partial x_i} + \sum_{j=1}^n \phi_j \frac{\partial}{\partial u_j}$$

si \mathbf{v} a pour composantes $(\xi_1, \dots, \xi_d, \phi_1, \dots, \phi_n)$.

Pour un groupe local de transformations associé à un groupe de Lie G , le théorème d'inversion locale montre que pour une fonction régulière $\mathbf{x} \mapsto \mathbf{u}(\mathbf{x})$, et pour $g \in G$ voisin de e , il existe localement une fonction $\tilde{\mathbf{x}} \mapsto \tilde{\mathbf{u}}(\tilde{\mathbf{x}})$ telle que

$$g \cdot (\mathbf{x}, \mathbf{u}(\mathbf{x})) = (\tilde{\mathbf{x}}, \tilde{\mathbf{u}}(\tilde{\mathbf{x}})).$$

Définition III.2

Considérons une EDP abstraite

$$(III.1) \quad F(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) = 0,$$

dans laquelle

$$\begin{aligned} \mathbf{u} : \mathbb{R}^d &\rightarrow \mathbb{R}^n \\ \mathbf{x} &\mapsto \mathbf{u}(\mathbf{x}) \end{aligned}$$

désigne la fonction inconnue, et pour tout $k \leq m$ (un entier au moins égal à 1), $d^k u$ est la différentielle d'ordre k de \mathbf{u} . On supposera la fonction

$$F : \Omega \times U \times \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n) \times \dots \times \mathcal{L}_m^s(\mathbb{R}^d; \mathbb{R}^n) \rightarrow \mathbb{R}^n$$

de classe \mathcal{C}^∞ (voire analytique).

Un groupe de symétrie de (III.1) est un groupe local de transformations agissant sur $\Omega \times U$ et laissant invariant l'ensemble des solutions régulières, c'est-à-dire que pour toute solution régulière $\mathbf{u} : \Omega \rightarrow U$ de (III.1), toute fonction \tilde{u} obtenue comme ci-dessus est aussi solution de (III.1).

Exemples

Invariance par translation : Une EDP à coefficients indépendants d'une variable x_i est invariante par translation $x_i \mapsto x_i + a$, quel que soit $a \in \mathbb{R}$. Ainsi \mathbb{R} , muni de l'addition, définit un groupe de symétrie de générateur infinitésimal ∂_{x_i} . Son action est simplement donnée par

$$a \cdot (\mathbf{x}, \mathbf{u}) = (x_1, \dots, x_i + a, \dots, x_d, \mathbf{u}).$$

Invariance par changement d'échelle : Pour l'équation de KdV

$$(III.2) \quad \partial_t u + \partial_{xxx}^3 u = u \partial_x u,$$

l'action de \mathbb{R}^{+*} donnée par

$$\lambda \cdot (t, x, u) = (\lambda^3 t, \lambda x, \lambda^{-2} u),$$

définit un groupe de symétrie de générateur infinitésimal

$$x \partial_x + 3t \partial_t - 2u \partial_u.$$

Dans le cas de (NLS) algébrique

$$(III.3) \quad i \partial_t u + \Delta u = \mu |u|^{p-1} u,$$

on trouve comme générateur de l'invariance par changement d'échelle

$$x \partial_x + 2t \partial_t - 2/(p-1) u \partial_u,$$

tandis que pour (NLW) algébrique

$$(III.4) \quad \partial_{tt}^2 u - c^2 \Delta u = \mu |u|^{p-1} u,$$

c'est

$$x \partial_x + t \partial_t - 2/(p-1) u \partial_u.$$

Invariance par changement de référentiel : L'équation (III.2) admet aussi un groupe de symétrie pour l'action de $(\mathbb{R}, +)$, définie par

$$a \cdot (t, x, u) = (t, x + at, u - a),$$

et de générateur infinitésimal $t\partial_x - \partial_u$.

Invariance par rotation : Le laplacien étant invariant par rotation, les équations de NLS et NLW le sont également. En dimension $d = 2$ par exemple, cela se traduit par l'invariance sous l'action de $(\mathbb{R}, +)$ définie par :

$$a \cdot (t, x, y, u) = (t, x \cos a - y \sin a, x \sin a + y \cos a, u),$$

de générateur infinitésimal $-y\partial_x + x\partial_y$.

Invariance par rotation de phase : L'équation (III.3) est invariante par l'action de \mathbb{R} définie par

$$\theta \cdot u = e^{i\theta} u,$$

dont le générateur infinitésimal est $iu\partial_u$.

Remarque III.1

Attention, diverses autres symétries peuvent exister, qui ne rentrent pas dans le cadre des symétries à un paramètre. C'est le cas notamment de l'inversion du sens du temps, $t \mapsto -t$ (avec $u \mapsto \bar{u}$ pour NLS), et de symétries plus compliquées (conforme, pseudo-conforme, Lorentz, cf [14, pp. 114-115]).

Il existe une méthode générale pour calculer *tous* les groupes de symétrie à un paramètre d'une EDP donnée. L'idée de départ est simple : elle consiste à considérer l'EDP comme un objet géométrique dans l'espace $\mathcal{E}^{d,n,m} := \mathbb{R}^d \times \mathbb{R}^n \times \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n) \times \dots \times \mathcal{L}_m^s(\mathbb{R}^d; \mathbb{R}^n)$. Un rapide calcul de la dimension de cet espace montre que la mise en œuvre de la méthode devient vite effroyablement technique lorsque d , n ou m augmentent. Mais par exemple pour l'équation de KdV (III.2), où $d = 1$, $n = 1$, $m = 3$, elle permet de montrer assez facilement que les seuls groupes de symétrie sont ceux que nous avons mis en évidence (invariance par translation en t et en x , invariance par changement d'échelle et invariance par changement de référentiel).

Avant d'énoncer le théorème général, voyons le cas d'une équation différentielle ordinaire autonome

$$(III.5) \quad \frac{d}{dt} \mathbf{u} = \mathbf{f}(\mathbf{u}),$$

et d'un champ

$$\mathbf{v} = \sum_{j=1}^n \phi_j \frac{\partial}{\partial u_j}$$

où les ϕ_j sont fonction de \mathbf{u} seulement. L'équation (III.5) est invariante par le groupe de symétries de générateur infinitésimal \mathbf{v} si pour toute solution \mathbf{u} , pour tout $g \in G$, la fonction $\tilde{\mathbf{u}} : t \mapsto \tilde{\mathbf{u}}(t) := g \cdot \mathbf{u}(t)$ est encore solution. Par dérivation par rapport à g on en déduit

$$(III.6) \quad \frac{d}{dt} \mathbf{v}(\mathbf{u}) = d\mathbf{f}(\mathbf{u}) \cdot \mathbf{v}(\mathbf{u}).$$

Il est d'usage, en regardant \mathbf{v} comme un opérateur différentiel, de noter le second membre de cette égalité comme $\mathbf{v}(\mathbf{f})(\mathbf{u})$, de sorte que $\mathbf{v}(\mathbf{f})$ désigne la fonction

$$\mathbf{v}(\mathbf{f}) : \mathbf{u} \mapsto \sum_{j=1}^n \phi_j(\mathbf{u}) \frac{\partial \mathbf{f}}{\partial u_j}(\mathbf{u}).$$

(Ainsi, au lieu de $\mathbf{v}(\mathbf{u})$, il faudrait noter $\mathbf{v}(\text{Id})(\mathbf{u})$: on identifie souvent les deux notations.) Quant au membre de gauche de (III.6), sa k -ième composante est

$$\sum_{j=1}^n \frac{\partial \phi_k}{\partial u_j}(\mathbf{u}) u'_j,$$

où $u'_j = du_j/dt$. Ainsi, une condition nécessaire pour l'invariance de (III.5) s'écrit de façon purement algébrique

$$(u'_k - f_k(\mathbf{u}) = 0, k = 1, \dots, n) \Rightarrow \left(\sum_{j=1}^n \frac{\partial \phi_k}{\partial u_j}(\mathbf{u}) u'_j - \mathbf{v}(f_k)(\mathbf{u}) = 0, k = 1, \dots, n \right).$$

Théorème III.1

Soit une EDP de la forme (III.1), où F est une fonction analytique de tous ses arguments. On suppose en outre qu'elle est de rang maximal (c'est-à-dire que la différentielle de F est de rang n en tout point de $\mathcal{E}^{d,n,m}$ où F s'annule), et qu'en tout point (d'un ouvert) de $\mathcal{E}^{d,n,m}$ il existe au moins une direction non caractéristique. Un champ de vecteurs

$$\mathbf{v} = \sum_{i=1}^d \xi_i \frac{\partial}{\partial x_i} + \sum_{j=1}^n \phi_j \frac{\partial}{\partial u_j}$$

est le générateur infinitésimal d'un groupe de symétrie pour (III.1) si et seulement si, quel que soit $(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) \in \mathcal{E}^{d,n,m}$,

$$F(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0 \implies \text{pr}^{(m)}\mathbf{v}(F)(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0,$$

où $\text{pr}^{(m)}\mathbf{v}$ désigne le champ de vecteurs prolongé à l'ordre m , défini par :

$$\text{pr}^{(m)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^n \sum_{1 \leq |\alpha| \leq m} \phi_j^\alpha \frac{\partial}{\partial u_{j,\alpha}},$$

avec

$$\phi_j^\alpha(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m\mathbf{u}) = D^\alpha \left(\phi_j - \sum_{i=1}^d \xi_i \frac{\partial u_j}{\partial x_i} \right) + \sum_{i=1}^d \xi_i \frac{\partial u_{j,\alpha}}{\partial x_i}.$$

Ci-dessus, D^α désigne la dérivée totale pour le multi-indice α , lorsque les ξ_i et ϕ_j sont évaluées en $(\mathbf{x}, \mathbf{u}(\mathbf{x}))$, et $u_{j,\alpha} = \partial^\alpha u_j$.

Exemples d'application

- La condition obtenue plus haut est un cas particulier du théorème, avec $d = 1$, $m = 1$, $\xi_i \equiv 0$,

$$F(\mathbf{u}, \mathbf{u}') = \mathbf{u}' - \mathbf{f}(\mathbf{u}), \quad \mathbf{v}(F_k) = -\mathbf{v}(f_k),$$

$$\text{pr}^{(1)}\mathbf{v} = \mathbf{v} + \sum_{j=1}^n \phi_j^1 \frac{\partial}{\partial u_j'} , \quad \phi_j^1 := D_t \phi_j = \sum_{\ell=1}^n \frac{\partial \phi_j}{\partial u_\ell}(\mathbf{u}) u_\ell' ,$$

$$\text{pr}^{(1)}\mathbf{v}(F_k) = \phi_k^1 - \mathbf{v}(f_k) .$$

- Pour l'équation de KdV (III.2), on a $d = 2$ (le temps t comptant pour une dimension et l'espace x pour une autre), $n = 1$, $m = 3$,

$$F(u, u_t, u_x, u_{xxx}) = u_t + u_{xxx} - uu_x .$$

Un champ

$$\mathbf{v} = \xi \partial_x + \tau \partial_t + \phi \partial_u$$

est le générateur infinitésimal d'un groupe de symétrie de (III.2) si

$$(III.7) \quad \phi^t + \phi^{xxx} - \phi u_x - u \phi^x = 0$$

pour tout (u, u_t, u_x, u_{xxx}) tel que $u_t + u_{xxx} - uu_x = 0$, avec

$$\phi^t := D_t(\phi - \xi u_x - \tau u_t) + \xi u_{xt} + \tau u_{tt} = \phi_t - u_x \xi_t + (\phi_u - \tau_t) u_t - u_x u_t \xi_u - \tau_u (u_t)^2 ,$$

$$\phi^x := D_x(\phi - \xi u_x - \tau u_t) + \xi u_{xx} + \tau u_{tx} = \phi_x - u_t \tau_x + (\phi_u - \xi_x) u_x - u_x u_t \tau_u - \xi_u (u_x)^2 ,$$

$$\phi^{xxx} := D_x^3(\phi - \xi u_x - \tau u_t) + \xi u_{xxx} + \tau u_{xxt} =$$

$$D_x^3 \phi - u_x D_x^3 \xi - u_t D_x^3 \tau - 3u_{xx} D_x^2 \xi - 3u_{xxx} D_x \xi - 3u_{tx} D_x^2 \tau - 3u_{txx} D_x \tau .$$

En remplaçant u_t par $uu_x - u_{xxx}$ (ainsi que u_{tx} par $(uu_x - u_{xxx})_x$, etc.) on peut tout écrire avec les « variables dépendantes » u, u_x, u_{xx} et u_{xxx} , et en collectant les différents facteurs dans (III.7), on peut montrer successivement (cf [12, p. 126]) que

1. le coefficient τ ne dépend que de t ,
2. le coefficient ξ ne dépend que de (t, x) ,
3. ils sont reliés par l'équation $\tau_t = 3\xi_x$,
4. le coefficient ϕ est linéaire en u et vérifie l'équation

$$\phi_t + \phi_{xxx} - u \phi_x = 0 ,$$

5. et qu'enfin

$$-\xi_t + (\phi_u - \tau_t) - (\phi_u - \xi_x) - \phi = 0 .$$

On en déduit qu'il existe quatre constantes $c_i, i = 1, \dots, 4$, tels que

$$\xi = c_1 + c_3 t + c_4 x , \quad \tau = c_2 + 3c_4 t , \quad \phi = -c_3 - 2c_4 u ,$$

c'est-à-dire que

$$\mathbf{v} = c_1 \partial_x + c_2 \partial_t + c_3 (t \partial_x - \partial_u) + c_4 (x \partial_x + 3t \partial_t - 2u \partial_u) .$$

Remarque III.2

Une façon équivalente d'écrire le champ prolongé est

$$\text{pr}^{(m)}\mathbf{v} = \text{pr}^{(m)}\mathbf{v}_Q + \sum_{i=1}^d \xi_i D_i ,$$

où

$$\mathbf{v}_Q := \sum_{j=1}^n Q_j \frac{\partial}{\partial u_j}, \quad Q_j := \phi_j - \sum_{i=1}^d \xi_i \frac{\partial u_j}{\partial x_i},$$

et D_i est la dérivée totale par rapport à x_i , qui s'écrit explicitement

$$D_i := \frac{\partial}{\partial x_i} + \sum_{j=1}^n \sum_{\alpha} \frac{\partial u_{j,\alpha}}{\partial x_i} \frac{\partial}{\partial u_{j,\alpha}}.$$

Symétries généralisées L'exemple de l'équation de KdV laisse entendre que les groupes de symétries au sens de la définition III.2 sont assez peu nombreux. C'est d'autant plus étonnant que cette équation est connue pour admettre une infinité de lois de conservation, et que (par le théorème de Noether que l'on verra plus loin), les lois de conservation sont associées à des symétries. Ceci nous incite à admettre des symétries généralisées, pour lequel les coefficients du champ \mathbf{v} sont autorisés à dépendre non seulement de (\mathbf{x}, \mathbf{u}) , mais aussi des dérivées de \mathbf{u} en nombre fini (comme c'est d'ailleurs le cas pour le champ \mathbf{v}_Q de la remarque ci-dessus).

Définition III.3

À un champ de vecteurs généralisé

$$\mathbf{v} = \sum_{i=1}^d \xi_i \frac{\partial}{\partial x_i} + \sum_{j=1}^n \phi_j \frac{\partial}{\partial u_j}$$

on associe le champ dit évolutionnaire

$$\mathbf{v}_Q = \sum_{j=1}^n Q_j \frac{\partial}{\partial u_j}, \quad Q_j := \phi_j - \sum_{i=1}^d \xi_i \frac{\partial u_j}{\partial x_i}.$$

On définit leurs prolongements par

$$\text{pr} \mathbf{v} = \text{pr} \mathbf{v}_Q + \sum_{i=1}^d \xi_i D_i, \quad \text{pr} \mathbf{v}_Q = \sum_{j=1}^n \sum_{\alpha} (D^\alpha Q_j) \frac{\partial}{\partial u_{j,\alpha}},$$

où $D^0 = \text{id}$.

Théorème III.2

Sous les hypothèses du théorème III.1, un champ de vecteurs généralisé \mathbf{v} est le générateur d'un groupe de symétrie généralisé pour (III.1) si et seulement si, quel que soit $(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)})$,

$$F(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0 \implies \text{pr} \mathbf{v}(F)(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0,$$

ou de façon équivalente,

$$F(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0 \implies \text{pr} \mathbf{v}_Q(F)(\mathbf{x}, \mathbf{u}, \dots, \mathbf{u}^{(m)}) = 0.$$

On voit qu'il « suffit » de considérer le représentant évolutionnaire \mathbf{v}_Q de \mathbf{v} . Cela vaut aussi pour les champs de vecteurs standard : par exemple, le champ $\mathbf{v} = -\partial_x$ a comme représentant évolutionnaire $\mathbf{v}_Q = u_x \partial_u$. Cependant, il s'avère que la notion de symétrie généralisée n'a d'intérêt que pour les EDP, c'est-à-dire pour $d \geq 2$. En effet, si \mathbf{v} est un champ de vecteurs généralisé, son représentant évolutionnaire \mathbf{v}_Q correspond à une symétrie généralisée pour une EDO

$$\frac{d\mathbf{u}}{dt} = \mathbf{f}(t, \mathbf{u})$$

si et seulement le champ de vecteurs standard

$$\sum_{j=1}^n \tilde{Q}_j \frac{\partial}{\partial u_j},$$

où \tilde{Q} est obtenu à partir de Q en substituant $\mathbf{f}(t, \mathbf{u})$ à \mathbf{u}' , etc., correspond à une symétrie ordinaire.

Pour les EDP d'évolution, le théorème III.2 peut se reformuler en terme de *crochets de Lie*.

Définition III.4

Soient des champs de vecteurs (standard)

$$\mathbf{v} = \sum_{j=1}^n \phi_j \frac{\partial}{\partial u_j}, \quad \mathbf{w} = \sum_{j=1}^n \psi_j \frac{\partial}{\partial u_j},$$

où les ϕ_j et ψ_j sont fonction de \mathbf{u} seulement. Leur crochet de Lie est le champ de vecteurs

$$[\mathbf{v}, \mathbf{w}] = \sum_{j=1}^n \sum_{i=1}^n \left(\phi_i \frac{\partial \psi_j}{\partial u_i} - \psi_i \frac{\partial \phi_j}{\partial u_i} \right) \frac{\partial}{\partial u_j}.$$

Avec la notation introduite plus haut, on peut aussi écrire

$$[\mathbf{v}, \mathbf{w}] = \sum_{j=1}^n (\mathbf{v}(\psi_j) - \mathbf{w}(\phi_j)) \frac{\partial}{\partial u_j},$$

et l'on observe que pour toute fonction régulière $f : U \rightarrow \mathbb{R}$,

$$[\mathbf{v}, \mathbf{w}](f) = \mathbf{v}(\mathbf{w}(f)) - \mathbf{w}(\mathbf{v}(f)).$$

Pour des champs de vecteurs généralisés \mathbf{v} et \mathbf{w} , le crochet de Lie est le champ de vecteurs généralisé $[\mathbf{v}, \mathbf{w}]$ tel que pour toute fonction régulière $f : \Omega \times U \times \mathcal{L}(\mathbb{R}^d; \mathbb{R}^n) \times \dots \times \mathcal{L}_m^s(\mathbb{R}^d; \mathbb{R}^n) \rightarrow \mathbb{R}$ (avec m assez grand),

$$\text{pr}[\mathbf{v}, \mathbf{w}](f) = \text{pr}\mathbf{v}(\text{pr}\mathbf{w}(f)) - \text{pr}\mathbf{w}(\text{pr}\mathbf{v}(f)).$$

Si $\mathbf{v} = \mathbf{v}_Q$ et $\mathbf{w} = \mathbf{v}_R$, alors $[\mathbf{v}, \mathbf{w}] = \mathbf{v}_S$ avec

$$S_j = \text{pr}\mathbf{v}_Q(R_j) - \text{pr}\mathbf{v}_R(Q_j).$$

Théorème III.3

Soit une EDP d'évolution

$$(III.8) \quad \frac{\partial \mathbf{u}}{\partial t} = P(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}),$$

où $P : \mathcal{E}^{d,n,m} \rightarrow \mathbb{R}^n$ est régulière et d désigne la différentiation par rapport à \mathbf{x} seulement. Soit le champ de vecteurs généralisé

$$\mathbf{v}_P = \sum_{j=1}^n P_j \frac{\partial}{\partial u_j},$$

où (P_1, \dots, P_n) sont les composantes de P . Un autre champ de vecteurs généralisé

$$\mathbf{v}_Q = \sum_{j=1}^n Q_j \frac{\partial}{\partial u_j},$$

où les Q_j dépendent de $(t, \mathbf{x}, \mathbf{u})$ et des dérivées de \mathbf{u} , est le générateur infinitésimal d'un groupe de symétries généralisées de (III.8) si et seulement si

$$\frac{\partial \mathbf{v}_Q}{\partial t} + [\mathbf{v}_P, \mathbf{v}_Q] = 0.$$

Démonstration : D'après le théorème III.2, une condition nécessaire et suffisante pour que \mathbf{v}_Q définisse une symétrie généralisée est

$$D_t Q_j = \text{pr} \mathbf{v}_Q(P_j), \quad j = 1, \dots, n.$$

En effet, ici $F(t, x, \mathbf{x}, \mathbf{u}, \mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m)}) = \mathbf{u}_t - P(x, \mathbf{x}, \mathbf{u}, \mathbf{u}^{(1)}, \dots, \mathbf{u}^{(m)})$ et (puisque P ne dépend pas de t) il n'y a pas d'autre multi-indice avec un premier coefficient non nul que $\alpha = (1, 0, \dots, 0)$ dans

$$\sum_{j=1}^n \sum_{\alpha} (D^{\alpha} Q_j) \frac{\partial F}{\partial u_{j,\alpha}}.$$

Or par définition de la dérivée totale,

$$D_t Q_j = \frac{\partial Q_j}{\partial t} + \sum_{k=1}^n \sum_{\alpha} \frac{\partial u_{k,\alpha}}{\partial t} \frac{\partial Q_j}{\partial u_{k,\alpha}} = \frac{\partial Q_j}{\partial t} + \sum_{k=1}^n \sum_{\alpha} (D^{\alpha} P)(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) \frac{\partial Q_j}{\partial u_{k,\alpha}}$$

pour \mathbf{u} solution de (III.8), et donc

$$D_t Q_j = \partial_t Q_j + \text{pr} \mathbf{v}_P(Q_j).$$

■

Remarque III.3

Puisque le crochet de Lie est antisymétrique, on voit en particulier que \mathbf{v}_P est le générateur infinitésimal d'un groupe de symétries généralisées de (III.8). Cependant, cette symétrie n'est rien d'autre que l'invariance par translation en t ! (dû au fait que P

ne dépend pas de t). En effet, le champ $-\partial_t$ admet comme représentant évolutionnaire

$$\sum_{j=1}^n u_{j,t} \frac{\partial}{\partial u_j},$$

qui coïncide avec \mathbf{v}_P sur l'ensemble des solutions de (III.8).

2 Lois de conservation

Définition III.5

Une loi de conservation (locale) pour une EDP abstraite de la forme (III.1) est une EDP « sous forme divergence »

$$D_1 e_1 + \cdots + D_d e_d = 0,$$

vérifiée par toutes les solutions régulières de (III.1), les D_i désignant (comme plus haut) les dérivées totales. On notera aussi $\text{Dive} = 0$, où \mathbf{e} est la fonction à valeurs dans \mathbb{R}^d de composantes e_1, \dots, e_d . En particulier, pour une équation d'évolution de la forme (III.8), une loi de conservation s'écrit

$$D_t e + \text{Div} \mathbf{q} = 0,$$

où Div désigne la divergence totale dans les variables spatiales. On dit que e est une densité conservée et que $\mathbf{q} = (q_1, \dots, q_d)$ est son flux. Par suite, la quantité

$$E(t) := \int_{\Omega} e(t, \mathbf{x}, \mathbf{u}(t, \mathbf{x}), \dots) d\mathbf{x}$$

est indépendante de t lorsque \mathbf{u} est une solution régulière de (III.1) tendant vers 0 ainsi que toutes ses dérivées lorsque \mathbf{x} s'approche de $\partial\Omega$.

Remarque III.4

Pour une EDO, une densité conservée n'est rien d'autre qu'une intégrale première.

Exemples

- L'équation (III.2) est elle-même une loi de conservation

$$D_t u + D_x (\partial_{xx}^2 u - \tfrac{1}{2} u^2) = 0.$$

Il en est de même pour l'équation de KdV généralisée :

$$D_t u + D_x (\partial_{xx}^2 u - f(u)) = 0.$$

Cette équation admet aussi comme loi de conservation

$$D_t (u^2) + D_x \left(u \partial_{xx}^2 u - \tfrac{1}{2} (\partial_x u)^2 - F(u) \right) = 0,$$

où F est une primitive de $u \mapsto u f'(u)$.

- L'équation (III.3) admet (notamment) comme lois de conservation

$$D_t |u|^2 + \operatorname{Div}(\operatorname{Im}(\bar{u} \nabla u)) = 0,$$

dont on verra qu'elle est liée à l'invariance par rotation de phase,

$$D_t(\operatorname{Im}(\bar{u} \partial_k u)) + \operatorname{Div} \operatorname{Re}(\partial_k u \nabla \bar{u}) - D_k \left(\frac{1}{4} \Delta |u|^2 \right) = 0, \quad k = 1, \dots, d,$$

liées à l'invariance par translation en espace, et

$$D_t \left(\frac{1}{2} |\nabla u|^2 + \frac{2\mu}{p+1} |u|^{p+1} \right) + \sum_{k=1}^d D_k \left(\frac{1}{2} \operatorname{Im}(\bar{\partial}_{k\ell}^2 u \partial_\ell u) + \mu |u|^{p-1} \operatorname{Im}(\bar{u} \partial_k u) \right) = 0,$$

liée à l'invariance par translation en temps.

- L'équation (III.4) admet comme lois de conservation

$$D_t \left(\frac{1}{2} |\partial_t u|^2 + \frac{1}{2} c^2 |\nabla u|^2 + \frac{\mu}{p+1} |u|^{p+1} \right) - c^2 \sum_{k=1}^d D_k \operatorname{Re}(\partial_t u \bar{\partial}_k u) = 0,$$

$$D_t(\operatorname{Re}(\partial_t \bar{\partial}_k u)) - \operatorname{Div} \operatorname{Re}(\bar{\partial}_k u \nabla u) + D_k \left(\frac{1}{2} c^2 |\nabla u|^2 - \frac{1}{2} |\partial_t u|^2 + \frac{2\mu}{p+1} |u|^{p+1} \right) = 0,$$

pour $k = 1, \dots, d$.

3 Formalisme Hamiltonien et théorème de Noether

3.1 EDO

Considérons un système d'équations différentielles ordinaires dans $\mathbb{R}^n \times \mathbb{R}^n$, de la forme

$$(III.9) \quad \begin{cases} \frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}(\mathbf{q}, \mathbf{p}), \\ \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}(\mathbf{q}, \mathbf{p}), \end{cases}$$

où le *Hamiltonien* $H : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ sera supposé de classe \mathcal{C}^∞ . Les solutions de (III.9) décrivent les courbes intégrales du *champ de vecteurs hamiltonien*

$$\widehat{\mathbf{v}}_H := \sum_{j=1}^n \frac{\partial H}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j} \frac{\partial}{\partial p_j}.$$

Plus généralement, à toute fonction $E : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ régulière, on peut associer le champ de vecteur hamiltonien

$$\widehat{\mathbf{v}}_E := \sum_{j=1}^n \frac{\partial E}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial E}{\partial q_j} \frac{\partial}{\partial p_j}.$$

Il s'avère que les intégrales premières de (III.9) sont intimement liées aux groupes de symétrie engendrés par les champs de vecteurs hamiltoniens. Tout d'abord, d'après le théorème III.3, un champ $\widehat{\mathbf{v}}_E$ est le générateur infinitésimal d'un groupe de symétries à un paramètre pour (III.9) si et seulement si

$$\frac{\partial \widehat{\mathbf{v}}_E}{\partial t} + [\widehat{\mathbf{v}}_H, \widehat{\mathbf{v}}_E] = 0.$$

Ensuite, le crochet de Lie $[\widehat{\mathbf{v}}_H, \widehat{\mathbf{v}}_E]$ est relié au *crochet de Poisson* des fonctions H et E .

Définition III.6

Le crochet de Poisson (associé à forme symplectique $\sum_j dq_j \wedge dp_j$) de deux fonctions régulières $E : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, $K : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$, est défini par

$$\{E, K\} = \sum_{j=1}^n \frac{\partial K}{\partial p_j} \frac{\partial E}{\partial q_j} - \frac{\partial K}{\partial q_j} \frac{\partial E}{\partial p_j}.$$

On constate ainsi que quelles que soient les fonctions régulières E et f ,

$$\widehat{\mathbf{v}}_E(f) = \{f, E\},$$

d'où

$$[\widehat{\mathbf{v}}_H, \widehat{\mathbf{v}}_E](f) = \widehat{\mathbf{v}}_H(\widehat{\mathbf{v}}_E(f)) - \widehat{\mathbf{v}}_E(\widehat{\mathbf{v}}_H(f)) = \{\{f, E\}, H\} - \{\{f, H\}, E\} = \{f, \{E, H\}\}$$

d'après l'*identité de Jacobi*, et par conséquent

$$[\widehat{\mathbf{v}}_H, \widehat{\mathbf{v}}_E] = \widehat{\mathbf{v}}_{\{E, H\}}.$$

Enfin, le calcul différentiel élémentaire montre que si E est une intégrale première de (III.9) alors

$$\frac{\partial E}{\partial t} + \{E, H\} = 0.$$

Proposition III.1

Si E est une intégrale première de (III.9) alors $\widehat{\mathbf{v}}_E$ est le générateur infinitésimal d'un groupe de symétries à un paramètre pour (III.9).

La réciproque est « presque » vraie (aux *fonctions Casimir* près, qui pour la forme symplectique $\sum_j dq_j \wedge dp_j$ sont simplement les fonctions constantes : par définition, une fonction Casimir C vérifie $\{C, f\} \equiv 0$ quelle que soit la fonction régulière f).

Théorème III.4 (Noether)

Soit $E : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ régulière. Le champ hamiltonien associé $\widehat{\mathbf{v}}_E$ est le générateur infinitésimal d'un groupe de symétries à un paramètre pour (III.9) si et seulement s'il existe une fonction $C : t \mapsto C(t) \in \mathbb{R}$ telle que $\tilde{E} := E - C$ soit une intégrale première de (III.9).

Démonstration : Si $\tilde{E} := E - C$ est une intégrale première de (III.9), alors $\widehat{\mathbf{v}}_{\tilde{E}} = \widehat{\mathbf{v}}_E$ est le générateur infinitésimal d'un groupe de symétries à un paramètre pour (III.9) d'après la proposition précédente. Réciproquement, si $\widehat{\mathbf{v}}_E$ est le générateur infinitésimal d'un groupe de symétries à un paramètre pour (III.9), on a

$$\frac{\partial \widehat{\mathbf{v}}_E}{\partial t} + [\widehat{\mathbf{v}}_H, \widehat{\mathbf{v}}_E] = 0,$$

c'est-à-dire

$$\widehat{\mathbf{v}}_{\frac{\partial E}{\partial t} + \{E, H\}} = 0,$$

d'où

$$\left\{ f, \frac{\partial E}{\partial t} + \{E, H\} \right\} = 0$$

quelle que soit f . Donc $\frac{\partial E}{\partial t} + \{E, H\}$ dépend seulement de t . En appelant C une primitive de cette fonction de t et $\tilde{E} := E - C$ on a

$$\frac{\partial \tilde{E}}{\partial t} + \{\tilde{E}, H\} = 0$$

et par conséquent \tilde{E} est une intégrale première de (III.9). ■

3.2 EDP

Le système d'EDO (III.9) est le « prototype » de système hamiltonien en dimension finie : le théorème de Darboux (cf par exemple [12, p. 405]) affirme même que c'est la *forme canonique*.

En dimension infinie, et plus spécialement pour des EDP d'évolution comme (III.8), on peut encore définir des structures hamiltoniennes, mais il n'y a pas de structure « canonique ». Ces structures reposent sur deux notions concernant les fonctionnelles : le gradient variationnel et le crochet de Poisson.

Définition III.7

Soit

$$\begin{aligned} H : \mathcal{E}^{d,n,m} &\rightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) &\mapsto H(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) \end{aligned}$$

une fonction régulière (\mathcal{C}^∞). On appelle fonctionnelle associée à H l'application qui à toute fonction \mathbf{u} régulière telle que $H(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u})$ soit intégrable sur \mathbb{R}^d associe

$$\mathcal{H}[\mathbf{u}] := \int_{\mathbb{R}^d} H(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) d\mathbf{x}.$$

Le gradient variationnel de \mathcal{H} est la fonction

$$\begin{aligned} \delta \mathcal{H} : \mathcal{E}^{d,n,m} &\rightarrow \mathbb{R}^n \\ (\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) &\mapsto \delta \mathcal{H}(\mathbf{x}, \mathbf{u}, d\mathbf{u}, \dots, d^m \mathbf{u}) \end{aligned}$$

caractérisée par

$$\frac{d\mathcal{H}[\mathbf{u} + \theta \mathbf{h}]}{d\theta} \Big|_{\theta=0} = \int \delta \mathcal{H}(\mathbf{x}, \mathbf{u}(\mathbf{x}), d\mathbf{u}(\mathbf{x}), \dots, d^m \mathbf{u}(\mathbf{x})) \cdot \mathbf{h}(\mathbf{x}) d\mathbf{x},$$

quel que soit $\mathbf{h} \in \mathcal{D}(\mathbb{R}^d; \mathbb{R}^n)$.

Grâce à des intégrations par parties, le gradient variationnel s'exprime à l'aide de l'opérateur d'Euler :

$$\delta \mathcal{H} = \begin{pmatrix} E_1 H \\ \vdots \\ E_n H \end{pmatrix},$$

$$E_j H = \sum_{\alpha} (-1)^{|\alpha|} D^{\alpha} \left(\frac{\partial H}{\partial u_{j,\alpha}} \right).$$

Exemple : si $d = 1$, $m = 1$, et $H = H(u, u_x)$,

$$EH = \frac{\partial H}{\partial u} - D_x \frac{\partial H}{\partial u_x}.$$

Dans la suite, pour toute fonction $\mathbf{u} : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^n$ assez régulière, on notera simplement $\delta \mathcal{H}[\mathbf{u}]$ la fonction

$$\begin{aligned} \delta \mathcal{H}[\mathbf{u}] : \mathbb{R} \times \mathbb{R}^d &\rightarrow \mathbb{R}^n \\ (t, \mathbf{x}) &\mapsto \delta \mathcal{H}(\mathbf{x}, \mathbf{u}(t, \mathbf{x}), d\mathbf{u}(t, \mathbf{x}), \dots, d^m \mathbf{u}(t, \mathbf{x})), \end{aligned}$$

la différentiation d ne portant que sur les variables spatiales (\mathbf{x}) .

Définition III.8

Une EDP d'évolution (III.8) admet une structure hamiltonienne s'il existe des applications régulières $J : \mathcal{E}^{d,n,m} \rightarrow \mathbb{R}^n$ et $H : \mathcal{E}^{d,n,m} \rightarrow \mathbb{R}$ telles que $P[\mathbf{u}] = \mathcal{J} \delta \mathcal{H}[\mathbf{u}]$, c'est-à-dire telles que l'EDP s'écrive :

$$(III.10) \quad \frac{\partial \mathbf{u}}{\partial t} = \mathcal{J} \delta \mathcal{H}[\mathbf{u}],$$

où \mathcal{H} est la fonctionnelle associée à H (cf Définition III.7), et \mathcal{J} est l'opérateur différentiel (éventuellement non-linéaire) défini par

$$(\mathcal{J} \mathbf{u})(\mathbf{x}) = J(\mathbf{x}, \mathbf{u}(\mathbf{x}), d\mathbf{u}(\mathbf{x}), \dots, d^m \mathbf{u}(\mathbf{x})),$$

auquel on demande de vérifier les deux propriétés suivantes :

Anti-symétrie : pour tout couple de fonctionnelles $(\mathcal{Q}, \mathcal{R})$,

$$\int \delta \mathcal{Q}[\mathbf{u}] \cdot \mathcal{J} \delta \mathcal{R}[\mathbf{u}] = - \int \delta \mathcal{R}[\mathbf{u}] \cdot \mathcal{J} \delta \mathcal{Q}[\mathbf{u}]$$

pour toute fonction \mathbf{u} régulière telle que les deux membres de l'égalité soient bien définis. Autrement dit, le crochet $\{, \}$ définissant la fonctionnelle $\{\mathcal{Q}, \mathcal{R}\}$ par

$$\{\mathcal{Q}, \mathcal{R}\}[\mathbf{u}] = \int \delta \mathcal{Q}[\mathbf{u}] \cdot \mathcal{J} \delta \mathcal{R}[\mathbf{u}]$$

est antisymétrique.

Identité de Jacobi : le crochet défini ci-dessus vérifie, pour tout triplet de fonctionnelles $(\mathcal{Q}, \mathcal{R}, \mathcal{S})$,

$$\{\mathcal{Q}, \{\mathcal{R}, \mathcal{S}\}\} + \{\mathcal{R}, \{\mathcal{S}, \mathcal{Q}\}\} + \{\mathcal{S}, \{\mathcal{Q}, \mathcal{R}\}\} = 0.$$

Ainsi $\{, \}$ est un crochet de Poisson.

Remarque III.5

Pour vérifier l'anti-symétrie de $\{, \}$, il suffit de vérifier que \mathcal{J} est un opérateur anti-symétrique sur $L^2(\mathbb{R}^d; \mathbb{R}^n)$, voir [12, Ex. 5.42, p. 385]. C'est en général assez facile. En revanche, il est beaucoup plus délicat de vérifier l'identité de Jacobi, sauf si J ne dépend en fait pas de \mathbf{u} ni des $d^k \mathbf{u}$, auquel cas l'identité de Jacobi est automatique dès que l'on a l'antisymétrie (voir [12, p. 438]).

Exemples de structures hamiltoniennes

KdV : L'équation de KdV généralisée

$$\partial_t u + \partial_{xxx}^3 u = \partial_x f(u)$$

s'écrit

$$\partial_t u = \mathcal{J} \delta \mathcal{H}[u]$$

avec $\mathcal{J} = D_x$ et

$$H(u, u_x) = \frac{1}{2}(u_x)^2 + F(u),$$

où F est une primitive de f . En effet,

$$\delta \mathcal{H}[u] = -\partial_{xx}^2 u + f(u).$$

Dans le cas particulier $f(u) = u^2/2$, une autre formulation hamiltonienne, moins « naturelle » est obtenue avec $H(u) = u^2/2$ et

$$\mathcal{J} = -D_x^3 + \frac{2}{3}u D_x + \frac{1}{3}u_x.$$

L'antisymétrie de \mathcal{J} se vérifie aisément par des intégrations par parties. Pour l'identité de Jacobi, voir [12, p. 442] (ceci demande d'assimiler quelques outils au préalable!).

NLS : L'équation de Schrödinger non-linéaire

$$i\partial_t u + \Delta u = f(u)$$

s'écrit aussi en identifiant \mathbb{C} à \mathbb{R}^2 comme

$$\partial_t u = J(-\Delta u + f(u)), \quad J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Par suite, si $f(u) = \nabla_u V(u)$, on obtient la formulation hamiltonienne :

$$\partial_t u = J \delta \mathcal{H}[u],$$

avec

$$H(u) = \frac{1}{2}(\|\partial_x u\|^2 + \|\partial_y u\|^2) + V(u),$$

où $\|\cdot\|$ désigne la norme euclidienne dans \mathbb{R}^2 .

Remarque III.6

Pour une EDP hamiltonienne (III.10) dans laquelle le hamiltonien ne dépend pas de t , la fonctionnelle \mathcal{H} est (au moins formellement) conservée le long des solutions. En effet, si $\mathbf{u} \in \mathcal{C}^1(0, T; \mathcal{D}(\mathbb{R}^d; \mathbb{R}^n))$,

$$\frac{d}{dt} \mathcal{H}[\mathbf{u}] = \int \delta \mathcal{H}[\mathbf{u}] \cdot \partial_t \mathbf{u} = \int \delta \mathcal{H}[\mathbf{u}] \cdot \mathcal{J} \delta \mathcal{H}[\mathbf{u}] = 0.$$

Ainsi l'invariance de (III.10) par translation en t est reliée à la conservation de \mathcal{H} : ceci est un cas particulier du théorème de Noether généralisé.

Étant donné l'opérateur \mathcal{J} , on peut associer à chaque fonctionnelle \mathcal{R} un *champ de vecteurs hamiltonien généralisé*

$$\widehat{\mathbf{v}}_{\mathcal{R}} := \sum_{j=1}^n Q_j \frac{\partial}{\partial u_j}, \quad Q_j[\mathbf{u}] := (\mathcal{J} \delta \mathcal{R}[\mathbf{u}])_j.$$

On peut ensuite faire agir le prolongement de ce champ $\widehat{\mathbf{v}}_{\mathcal{R}}$ sur les fonctionnelles \mathcal{P} en posant

$$\text{pr} \widehat{\mathbf{v}}_{\mathcal{R}}(\mathcal{P}) := \int \text{pr} \widehat{\mathbf{v}}_{\mathcal{R}}(P).$$

Lemme III.1

Si \mathcal{J} définit un crochet de Poisson $\{, \}$, on a

$$\text{pr} \widehat{\mathbf{v}}_{\mathcal{R}}(\mathcal{P}) = \{ \mathcal{P}, \mathcal{R} \}.$$

Démonstration : Ceci repose sur l'observation cruciale suivante, que l'on admettra dans le cas général (voir [12, p. 367]) : pour tout couple d'applications régulières $P, R : \mathcal{E}^{d,n,m} \rightarrow \mathbb{R}$, il existe des applications A_1, \dots, A_d telles que

$$\text{pr} \mathbf{v}_Q(P) = \sum_{j=1}^n Q_j E_j P + \sum_{k=1}^d D_k A_k.$$

Ci-dessus, E_j désigne l'opérateur d'Euler. Dans le cas particulier $d = 1, m = 1$, et $H = H(u, u_x)$, on voit facilement d'où vient la formule :

$$\text{pr} \mathbf{v}_Q(P) = Q \frac{\partial P}{\partial u} + (D_x Q) \frac{\partial P}{\partial u_x} = Q E P + D_x \left(Q \frac{\partial P}{\partial u_x} \right)$$

puisque

$$E P = \frac{\partial P}{\partial u} - D_x \frac{\partial P}{\partial u_x}.$$

Ainsi, les dérivées exactes ne contribuant pas à l'intégrale, on a

$$\text{pr} \widehat{\mathbf{v}}_{\mathcal{R}}(\mathcal{P}) = \int \text{pr} \mathbf{v}_Q(P) = \int \sum_{j=1}^n Q_j E_j P = \int \delta \mathcal{P} \cdot \mathcal{J} \delta \mathcal{R}$$

puisque $Q_j = (\mathcal{J} \delta \mathcal{R})_j$ et $E_j P = (\delta \mathcal{P})_j$. ■

Théorème III.5 (Noether généralisé)

Soit $R : \mathcal{E}^{d,n,m} \rightarrow \mathbb{R}$ régulière. Le champ hamiltonien généralisé $\widehat{\mathbf{v}}_{\mathcal{R}}$ est le générateur infinitésimal d'un groupe de symétries généralisées à un paramètre pour l'EDP hamiltonienne (III.10) si et seulement s'il existe $\mathcal{C} : t \mapsto \mathcal{C}(t)$ telle que $\mathcal{J} \delta \mathcal{C}(t) = 0$ quel que soit $t \in \mathbb{R}$ et $\widetilde{\mathcal{R}} := \mathcal{R} - \mathcal{C}$ soit conservée le long des solutions de (III.10).

Démonstration : La démonstration est calquée sur celle du théorème de Noether en dimension finie. D'après le théorème III.3, une condition nécessaire et suffisante pour que $\widehat{\mathbf{v}}_{\mathcal{R}}$ soit le générateur infinitésimal d'un groupe de symétries généralisées à un paramètre est

$$\frac{\partial \widehat{\mathbf{v}}_{\mathcal{R}}}{\partial t} + [\widehat{\mathbf{v}}_{\mathcal{H}}, \widehat{\mathbf{v}}_{\mathcal{R}}] = 0.$$

Or on vérifie (cf [12, p. 447]) à l'aide des définitions des crochet de Lie et de Poisson, du lemme III.1 et de l'identité de Jacobi que

$$[\widehat{\mathbf{v}}_{\mathcal{H}}, \widehat{\mathbf{v}}_{\mathcal{R}}] = \widehat{\mathbf{v}}_{\{\mathcal{R}, \mathcal{H}\}}.$$

Donc la condition s'écrit

$$\widehat{\mathbf{v}}_{\frac{\partial \mathcal{R}}{\partial t} + \{\mathcal{R}, \mathcal{H}\}} = 0,$$

d'où

$$\left\{ \mathcal{P}, \frac{\partial \mathcal{R}}{\partial t} + \{\mathcal{R}, \mathcal{H}\} \right\} = 0$$

quelle que soit la fonctionnelle \mathcal{P} . Ceci équivaut à ce que $\frac{\partial \mathcal{R}}{\partial t} + \{\mathcal{R}, \mathcal{H}\}$ soit une fonctionnelle $\widetilde{\mathcal{C}}(t)$ vérifiant $\mathcal{J} \delta \widetilde{\mathcal{C}}(t) = 0$. En appelant $\mathcal{C}(t)$ une fonctionnelle obtenue en intégrant $\widetilde{\mathcal{C}}$ par rapport à t , et $\widetilde{\mathcal{R}} := \mathcal{R} - \mathcal{C}$, on obtient la condition équivalente

$$\frac{\partial \widetilde{\mathcal{R}}}{\partial t} + \{\widetilde{\mathcal{R}}, \mathcal{H}\} = 0,$$

qui signifie précisément que $\widetilde{\mathcal{R}}$ est une quantité conservée le long des solutions de (III.10). ■

Exemples.

Invariance par translation en temps. Le théorème de Noether confirme le lien entre la conservation de l'énergie \mathcal{H} et l'invariance par translation en t . En effet, le groupe de symétrie associé à l'invariance par translation en t a comme représentant évolutionnaire

$$u_t \partial_u = (\mathcal{J} \delta \mathcal{H}[u]) \partial_u = \widehat{\mathbf{v}}_{\mathcal{H}}.$$

Invariance par translation en espace. Dans le cas monodimensionnel $d = 1, n = 1$, le groupe de symétrie associé à l'invariance par translation en x a comme représentant évolutionnaire $u_x \partial_u$. Lorsque $\mathcal{J} = D_x$ (comme pour KdV), il s'écrit comme un champ de vecteurs hamiltonien généralisé $\widehat{\mathbf{v}}_{\mathcal{R}}$ avec $\mathcal{R} = \frac{1}{2} \int u^2$. En effet, $\delta \mathcal{R}[u] = u$, et par conséquent $u_x = \mathcal{J} \delta \mathcal{R}[u]$. On voit ainsi que l'invariance par translation en x est liée à la conservation de la norme L^2 .

En dimension supérieure, pour chaque $k \in \{1, \dots, d\}$, le représentant évolutionnaire associé à l'invariance par translation en x_k est $\sum_{j=1}^n u_{j,x_k} \partial_{u_j}$. En particulier pour $n = 2$ et $\mathcal{J} = -i$ (comme c'est le cas pour NLS), on a $u_{x_k} = \mathcal{J} \delta \mathcal{R}_k[u]$ avec $R_k[u] = \text{Im}(\bar{u} u_{x_k})$.

Invariance par rotation de phase, (pour (NLS) algébrique par exemple). Le champ de vecteurs généralisé correspondant est $-iu\partial_u$, soit en coordonnées réelles, $u_2\partial_{u_1} - u_1\partial_{u_2} = \widehat{\mathbf{v}}_{\mathcal{R}}$ avec $\mathcal{R} = \frac{1}{2} \int \|u\|^2$, si $\mathcal{J} = -i$.

Autres invariances pour KdV. On a déjà vu deux formulations hamiltoniennes pour KdV. Dans la première, pour laquelle $\mathcal{J} = D_x$, les fonctionnelles Casimir (qui satisfont $\mathcal{J}\delta\mathcal{C} = 0$) sont proportionnelles à la masse $\mathcal{M} = \int u$. Cette dernière est donc conservée (ce qui se voit bien sûr directement sur la *forme conservative* de l'EDP). D'autre part, on a vu que KdV était invariante par le groupe de symétrie engendré par $t\partial_x - \partial_u$ (invariance par changement de référentiel), dont le représentant évolutionnaire est $-(1 + tu_x)\partial_u$. En observant que $1 + tu_x = D_x\delta\mathcal{R}[u]$ avec $\mathcal{R} = xu + \frac{1}{2}tu^2$, et en utilisant la conservation de la norme L^2 (associée à l'invariance par translation en x), on en déduit que le moment d'ordre 1 de toute solution u , $\int xu(t, x)dx$ est une fonction affine de t .

Quant à l'invariance par changement d'échelle, elle ne donne rien dans la formulation avec $\mathcal{J} = D_x$ (c'est-à-dire que pour le représentant évolutionnaire $Q\partial_u$ de $x\partial_x + 3t\partial_t - 2u\partial_u$ on ne peut pas écrire $Q[u] = D_x\delta\mathcal{R}[u]$), tandis qu'elle redonne le moment d'ordre 1 dans la formulation avec $\mathcal{J} = -D_x^3 + \frac{2}{3}uD_x + \frac{1}{3}u_x$.

Cependant, on peut obtenir de nouvelles lois de conservation en « jonglant » avec les deux structures hamiltoniennes. En effet, la partie « facile » du théorème de Noether permet (re)trouver une nouvelle symétrie, associée à $Q\partial_u$ avec

$$Q[u] = (-D_x^3 + \frac{2}{3}uD_x + \frac{1}{3}u_x) \left(-u_{xx} + \frac{1}{2}u^2\right),$$

(on rappelle que $-u_{xx} + \frac{1}{2}u^2 = \delta\mathcal{H}[u]$ avec $H(u, u_x) = \frac{1}{2}(u_x)^2 + \frac{1}{6}u^3$), et il s'avère que

$$Q[u] = D_x\delta\mathcal{R}[u]$$

pour une nouvelle fonctionnelle \mathcal{R} , qui est donc conservée. (Tous calculs faits, $Q[u] = u_{xxxxx} - \frac{5}{3}uu_{xxx} - \frac{10}{3}u_xu_{xx} + \frac{5}{6}u^2u_x$ et $R[u] = \frac{1}{2}u_{xx}^2 + \frac{5}{6}uu_x^2 + \frac{5}{72}u^4$.)

Chapitre IV

Solitons

On a vu au chapitre précédent que les principales équations dispersives étudiées dans ce cours étaient dotées de (nombreuses) structures hamiltoniennes, symétries, et lois de conservation. Elles ont un autre point commun remarquable : celui d'admettre des solutions exactes très particulières, que l'on appelle *solitons*. Le terme « soliton » a été inventé dans les années 1960, à partir d'« onde solitaire » (terme déjà employé par Boussinesq en 1871) et du suffixe -on qui rappelle celui des particules élémentaires car les solitons sont remarquablement stables aux interactions. L'objectif de ce chapitre est de donner quelques résultats concernant l'existence et la stabilité des solitons, en dimension 1 d'espace.

1 Existence des solitons

Une *onde solitaire* est une onde plane progressive « localisée », c'est-à-dire qu'elle a des limites en $-\infty$ et $+\infty$ et que ces limites sont les mêmes (généralement 0). Comme toutes les ondes planes progressives les ondes solitaires sont définies à l'aide de leur *profil* et de leur vitesse, qui est un scalaire en dimension 1 d'espace. Ainsi un soliton de profil U et de vitesse σ est une fonction de la forme

$$u(t, x) = U(x - \sigma t) \quad \text{avec} \quad \lim_{\xi \rightarrow \pm\infty} U(\xi) = U_\infty ,$$

cette limite étant atteinte « rapidement » (en général exponentiellement vite), si bien que toutes les dérivées de U tendent vers zéro à l'infini. Dans toute la suite on s'intéressera aux solitons pour lesquels $U_\infty = 0$. Pour l'équation de Schrödinger, les solitons relèvent d'une définition un peu plus générale, qui autorise la phase de la solution à se propager à une vitesse éventuellement différente de σ . Ainsi pour (NLS) algébrique

$$(IV.1) \quad i\partial_t u + \Delta u = \mu |u|^{p-1} u ,$$

un soliton est une solution de la forme

$$u(t, x) = e^{i(rx-st)} U(x - \sigma t) \quad \text{avec} \quad \lim_{\xi \rightarrow \pm\infty} U(\xi) = 0 ,$$

et U à valeurs réelles : ceci est possible pour $r = \sigma/2$, car on trouve l'équation de profil à coefficients réels

$$U'' + (s - r^2)U - \mu |U|^{p-1}U = 0 .$$

Une façon « élémentaire » de montrer l'existence d'un soliton est de calculer explicitement son profil : ceci est possible pour certaines non-linéarités, et demande une certaine habileté technique !

Exemples. Pour (NLS) cubique ($p = 3$) dans le cas focalisant ($\mu < 0$), on vérifie que

$$U(\xi) = \sqrt{\frac{2\alpha}{-\mu}} \operatorname{sech}(\xi\sqrt{\alpha})$$

fournit une famille de solitons (avec $\alpha = r^2 - s$). La fonction *sécante hyperbolique* est définie comme $\operatorname{sech}(x) = 1/\operatorname{ch}(x)$ et apparaît très souvent dans les formules explicites de solitons.

Considérons maintenant l'équation de Korteweg-de Vries avec une non-linéarité quadratique :

$$u_t + uu_x + u_{xxx} = 0.$$

(Attention le choix du signe devant la non-linéarité est opposé à celui du chapitre précédent : ce changement est inoffensif puisqu'il revient à changer u en $-u$.) Pour trouver des solutions explicites, une première idée (comme pour l'équation de Burgers $u_t + uu_x = 0$) consiste à chercher u comme la dérivée en x d'une nouvelle fonction inconnue p , de sorte que

$$p_t + \frac{1}{2}p_x^2 + p_{xxx} = 0.$$

La seconde idée (comme dans la transformation de Cole-Hopf pour l'équation de Burgers-Hopf $u_t + uu_x = u_{xx}$) est de chercher p sous la forme

$$p = a \frac{F_x}{F},$$

où a est un paramètre à déterminer. Après calcul, on constate en effet que pour $a = 12$, l'équation devant être satisfaite par F se « réduit » à l'équation quadratique

$$F(F_t + F_{xxx})_x - F_x(F_t + F_{xxx}) + 3(F_{xx}^2 - F_x F_{xxx}) = 0.$$

Bien que cette dernière équation semble à première vue plus compliquée que (KdV), on en trouve des solutions explicites en remarquant que l'équation d'Airy

$$F_t + F_{xxx} = 0$$

admet des solutions ondes progressives vérifiant aussi

$$F_{xx}^2 - F_x F_{xxx} = 0.$$

Ainsi, pour tout $\alpha \in \mathbb{R}$ et pour tout $s \in \mathbb{R}$, $F(t, x) = 1 + e^{-\alpha(x-s) + \alpha^3 t}$ est solution de l'équation quadratique ci-dessus, d'où

$$p(t, x) = 12(\log F)_x = 6\alpha \tanh\left(\frac{1}{2}(-\alpha(x-s) + \alpha^3 t)\right),$$

et la famille de solitons définie par

$$u(t, x) = 3\alpha^2 \operatorname{sech}^2\left(\frac{1}{2}(-\alpha(x-s) + \alpha^3 t)\right)$$

Une approche un peu plus moderne pour démontrer l'existence de solitons consiste à étudier le *portrait de phase* de l'équation de profil. Prenons par exemple l'équation de Korteweg-de Vries généralisée

$$(IV.2) \quad u_t + u_{xxx} = f(u)_x,$$

avec $f(0) = 0$. L'équation de profil

$$-\sigma U' + U''' = f(U)'$$

s'intègre à vue en

$$U'' = f(U) + \sigma U.$$

Cette équation différentielle du second ordre s'écrit dans le *plan de phase* $\{(U, U')\}$ comme le système du premier ordre

$$(IV.3) \quad \begin{cases} U' = V \\ V' = f(U) + \sigma U \end{cases}$$

Le fait que f s'annule en 0 assure que $(0, 0)$ est un point fixe. Les solitons, s'ils existent, correspondent à des *orbites homoclines* au point fixe $(0, 0)$. Ce dernier doit donc être un point selle, ce qui demande $f'(0) + \sigma > 0$, et il doit exister un autre point fixe $(u_*, 0)$ (à l'intérieur de la séparatrice consistée par l'orbite homocline) qui soit un centre, et donc tel que $f'(u_*) + \sigma < 0$. C'est le cas dans l'exemple vu précédemment : $f(u) = -u^2/2$, $\sigma > 0$, $u_* = 2\sigma$. On constate en outre que le système (IV.3) admet comme intégrale première

$$\frac{1}{2}V^2 - F(U) - \frac{1}{2}\sigma U^2,$$

où F est une primitive de f . Le point $(u_m, 0)$ où l'orbite homocline, si elle existe, coupe l'axe des U doit donc être tel que

$$F(u_m) - F(0) - \frac{1}{2}\sigma u_m^2 = 0.$$

Dans l'exemple, $F(U) = -\frac{1}{6}u^3$, et $u_m = 3\sigma$: on retrouve ainsi presque sans calcul la valeur de l'amplitude du soliton (dont on constate qu'elle est d'autant plus grande que le soliton est rapide).

Le fait que l'équation de profil admette une intégrale première, et soit même hamiltonien, n'est pas un hasard, comme le montre le résultat abstrait suivant.

Proposition IV.1 (*Benjamin*)

Considérons une EDP hamiltonienne en dimension 1 d'espace

$$\partial_t \mathbf{u} = \mathcal{J} \delta \mathcal{H}[\mathbf{u}],$$

avec $\mathcal{J} = D_x \mathbf{J}$, où \mathbf{J} est une matrice symétrique inversible telle qu'il existe une fonctionnelle (associée à l'invariance par translation en x) $\mathcal{J} : \mathbf{u} \mapsto \mathcal{J}[\mathbf{u}] = \int I(\mathbf{u})$ vérifiant

$$\mathbf{u} = -\mathbf{J} \delta \mathcal{J}[\mathbf{u}].$$

On suppose que \mathcal{H} ne dépend que de \mathbf{u} et \mathbf{u}_x . Alors

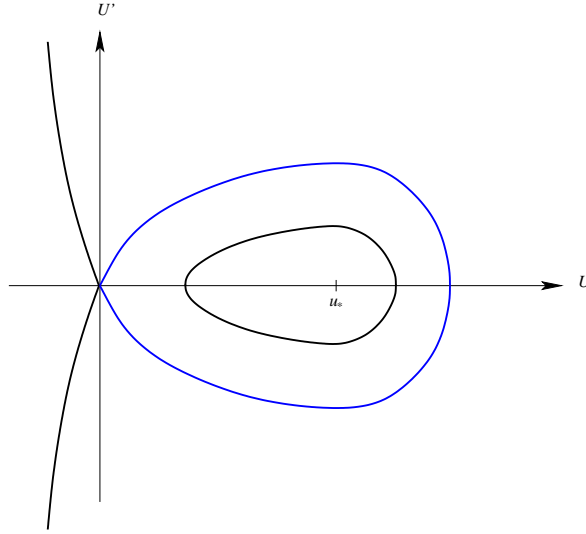


FIG. IV.1 – Portrait de phase présentant une orbite homocline

- les solutions de l'EDP vérifient une loi de conservation

$$I_t + S_x = 0,$$

où S est une fonction de \mathbf{u} , \mathbf{u}_x et \mathbf{u}_{xx} ;

- l'équation de profil des solitons de vitesse σ est l'équation d'Euler-Lagrange associée au lagrangien $H - \sigma I$. Le hamiltonien associé (et donc une intégrale première de l'équation de profil) étant donné par $S - \sigma I$.

Démonstration : Cherchons d'abord S . On doit avoir

$$S_x = -I_t = -\sum_{j=1}^n u_{j,t} \frac{\partial I}{\partial u_j} = -\sum_{j=1}^n u_{j,t} \frac{\partial I}{\partial u_j}.$$

Or si \mathbf{u} est solution de l'EDP,

$$u_{j,t} = \sum_{i=1}^n J_{ji} (E_i H)_x,$$

où E_i désigne l'opérateur d'Euler d'indice i : puisque \mathcal{H} ne dépend que de \mathbf{u} et \mathbf{u}_x ,

$$E_i H = -D_x \left(\frac{\partial H}{\partial u_{i,x}} \right) + \frac{\partial H}{\partial u_i}.$$

En utilisant la symétrie de \mathbf{J} et la définition de I , on en déduit

$$S_x = \sum_{i=1}^n u_i (E_i H)_x.$$

Par suite,

$$S = \sum_{i=1}^n u_i (E_i H) - H + \sum_{i=1}^n u_{i,x} \frac{\partial H}{\partial u_{i,x}}.$$

D'autre part, l'équation de profil s'écrit

$$-\sigma \mathbf{U}' = \mathbf{J} \delta \mathcal{H}[\mathbf{U}]',$$

soit encore

$$\mathbf{J}\delta(\mathcal{H} - \sigma\mathcal{I})[\mathbf{U}]' = 0.$$

Puisque \mathbf{J} est inversible on obtient par intégration (en supposant $\delta(\mathcal{H} - \sigma\mathcal{I})[0] = 0$)

$$\delta(\mathcal{H} - \sigma\mathcal{I})[\mathbf{U}] = 0,$$

c'est-à-dire

$$E_i(H - \sigma I) = 0, \quad i = 1, \dots, n.$$

Le hamiltonien associé à cette équation d'Euler-Lagrange est la transformée de Legendre du lagrangien $H - \sigma I$, c'est-à-dire (puisque I ne dépend pas de \mathbf{u}_x) la fonction

$$\sum_{i=1}^n u_{i,x} \frac{\partial H}{\partial u_{i,x}} - (H - \sigma I)$$

exprimée en variables u_i et $p_i = \frac{\partial H}{\partial u_{i,x}}$. Finalement, d'après l'équation de profil $\sigma E_i H = E_i I = \frac{\partial I}{\partial u_i}$,

$$\sum_{i=1}^n u_i(E_i H) = \sigma \sum_{i=1}^n u_i(E_i I) = 2\sigma I$$

puisque I est quadratique (par définition). ■

2 Stabilité des solitons

Commençons par préciser différentes notions de stabilité.

Définition IV.1

Considérons une EDP d'évolution

$$(IV.4) \quad \partial_t \mathbf{u} = P(\mathbf{u}, \mathbf{u}_x, \dots, \mathbf{u}_x^{(m)}),$$

et supposons qu'elle admette une solution stationnaire $\underline{\mathbf{u}}$ (c'est-à-dire que $P(\underline{\mathbf{u}}, \underline{\mathbf{u}}_x, \dots, \underline{\mathbf{u}}_x^{(m)}) = 0$) dans un espace fonctionnel X .

- On dit que $\underline{\mathbf{u}}$ est stable dans X si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour toute donnée initiale $\mathbf{u}_0 \in X$ telle que

$$\|\mathbf{u}_0 - \underline{\mathbf{u}}\|_X \leq \eta,$$

le problème de Cauchy associé à (IV.4) admet une solution valant \mathbf{u}_0 en $t = 0$, définie pour tout $t \in \mathbb{R}^+$, et vérifiant

$$\|\mathbf{u}(t) - \underline{\mathbf{u}}\|_X \leq \varepsilon.$$

- On dit que $\underline{\mathbf{u}}$ est asymptotiquement stable si elle est stable et s'il existe $\eta_0 > 0$ tel que, pour toute donnée initiale $\mathbf{u}_0 \in X$ telle que

$$\|\mathbf{u}_0 - \underline{\mathbf{u}}\|_X \leq \eta_0,$$

la solution $\mathbf{u}(t)$ tend vers $\underline{\mathbf{u}}$ lorsque t tend vers $+\infty$.

Même si un soliton n'est pas exactement une solution stationnaire, on peut toujours se ramener à une solution stationnaire, pour un soliton de vitesse σ , en faisant le changement de référentiel $(t, x) \mapsto (t, x - \sigma t)$. Cependant, puisque P ne dépend pas de x , (IV.4) est invariante par translation en x , et par conséquent tout profil \mathbf{U} de soliton est associé en fait à une infinité de solitons obtenus par translations :

$$\mathbf{u}_s(t, x) = \mathbf{U}(x - \sigma t + s) =: \mathbf{U}_s(x - \sigma t).$$

(Attention, on utilise ici l'indice s pour signifier une simple translation, et non une dérivation !) Pour cette raison, on doit assouplir la notion de stabilité.

Définition IV.2

Une onde progressive $\mathbf{u}(t, x) = \mathbf{U}(x - \sigma t)$ solution de (IV.4) est dite orbitalement stable si, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, pour toute donnée initiale $\mathbf{u}_0 \in X$ telle que

$$\|\mathbf{u}_0 - \mathbf{U}\|_X \leq \eta,$$

le problème de Cauchy associé à (IV.4) admet une solution valant \mathbf{u}_0 en $t = 0$, définie pour tout $t \in \mathbb{R}^+$, et vérifiant

$$\inf_{s \in \mathbb{R}} \|\mathbf{u}(t) - \mathbf{U}_s\|_X \leq \varepsilon.$$

On considère désormais une EDP hamiltonienne

$$(IV.5) \quad \partial_t \mathbf{u} = \mathcal{J} \delta \mathcal{H}[\mathbf{u}],$$

avec $\mathcal{J} = D_x \mathbf{J}$, où \mathbf{J} est une matrice symétrique inversible telle qu'il existe une fonctionnelle (associée à l'invariance par translation en x) $\mathcal{I} : \mathbf{u} \mapsto \mathcal{I}[\mathbf{u}] = \int I(\mathbf{u})$ vérifiant

$$\mathbf{u} = -\mathbf{J} \delta \mathcal{I}[\mathbf{u}].$$

On a vu dans la proposition IV.1 que les solitons $\mathbf{u}(t, x) = \mathbf{U}(x - \sigma t)$ correspondaient à des points critiques de la fonctionnelle $\mathcal{H} - \sigma \mathcal{I}$, leur équation de profil s'écrivant

$$(IV.6) \quad \delta(\mathcal{H} - \sigma \mathcal{I})[\mathbf{U}] = 0.$$

La fonctionnelle $\mathcal{H} - \sigma \mathcal{I}$ étant conservée le long des solutions de (IV.5), si jamais \mathbf{U} en est un minimum local (dans un espace approprié), on peut espérer l'utiliser comme *fonctionnelle de Lyapunov* pour démontrer la stabilité de \mathbf{u} , qui équivaut à la stabilité de \mathbf{U} comme solution stationnaire de l'équation (IV.5) réécrite dans le référentiel se déplaçant à la vitesse σ , c'est-à-dire

$$(IV.7) \quad \partial_t \mathbf{u} = \mathcal{J} \delta(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}].$$

Mais comme on l'a suggéré ci-dessus, on s'attend seulement à de la stabilité orbitale, à cause de l'invariance par translation en x . Et en effet, si l'on considère la *hessienne* de $\mathcal{H} - \sigma \mathcal{I}$, c'est-à-dire l'application $\mathbf{u} \mapsto$ l'opérateur différentiel $\text{Hess}(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}]$ défini par

$$\frac{d^2(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u} + \theta \mathbf{h}]}{d\theta^2} \Big|_{\theta=0} = \int \mathbf{h} \cdot \text{Hess}(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}] \mathbf{h},$$

on constate par dérivation de l'équation de profil que

$$\text{Hess } (\mathcal{H} - \sigma \mathcal{I})[\mathbf{U}]\mathbf{U}' = 0.$$

Par conséquent, \mathbf{U} ne peut pas être un minimum local strict de $\mathcal{H} - \sigma \mathcal{I}$. Au mieux, on peut espérer une minoration du type

$$(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}] - (\mathcal{H} - \sigma \mathcal{I})[\mathbf{U}] \geq \alpha \|\mathbf{u} - \mathbf{U}\|^2$$

avec $\alpha > 0$ dans un supplémentaire de la droite engendrée par \mathbf{U}' . Avec une telle minoration dans $(\mathbf{U}')^\perp$, on pourrait conclure à la stabilité orbitale de \mathbf{U} comme solution de (IV.7) en « éliminant » le « problème » lié à l'invariance par translation grâce au résultat suivant.

Lemme IV.1

Soit $\mathbf{U} \in H^m(\mathbb{R}; \mathbb{R}^n)$ avec $m \geq 1$. On note $\mathbf{U}_s : \xi \mapsto \mathbf{U}(\xi + s)$ ses translatés pour $s \in \mathbb{R}$, qui sont aussi dans $H^m(\mathbb{R}; \mathbb{R}^n)$. Il existe $\varepsilon_0 > 0$ et une application de classe \mathcal{C}^1

$$\tau : \mathcal{V} := \{\mathbf{V} \in H^m(\mathbb{R}; \mathbb{R}^n) ; \min_{s \in \mathbb{R}} \|\mathbf{V} - \mathbf{U}_s\|_{H^m} < \varepsilon_0\} \rightarrow \mathbb{R}$$

telle que, pour tout $\mathbf{V} \in \mathcal{V}$,

$$\int_{\mathbb{R}} \mathbf{V}_{\tau(\mathbf{V})} \cdot \mathbf{U}' = 0,$$

et pour tout $r \in \mathbb{R}$,

$$\tau(\mathbf{V}_r) = \tau(\mathbf{V}) - r.$$

Démonstration : On commence par construire τ au voisinage de \mathbf{U} , en appliquant le théorème des fonctions implicites à l'application $\mathcal{T} : (t, \mathbf{V}) \mapsto \int \mathbf{V}_t \cdot \mathbf{U}'$ au point $(0, \mathbf{U})$. En effet, \mathcal{T} est bien définie et de classe \mathcal{C}^1 sur $\mathbb{R} \times H^m(\mathbb{R}; \mathbb{R}^n)$ pour $m \geq 1$, et l'on a

$$\mathcal{T}(0, \mathbf{U}) = \int \mathbf{U} \cdot \mathbf{U}' = \left[\frac{1}{2} \|\mathbf{U}\|^2 \right]_{-\infty}^{+\infty} = 0, \quad \frac{\partial \mathcal{T}}{\partial t}(0, \mathbf{U}) = \int \|\mathbf{U}'\|^2 > 0.$$

On obtient donc par le théorème des fonctions implicites une application τ définie sur un voisinage $\mathcal{V}_0 := \{\mathbf{V} ; \|\mathbf{V} - \mathbf{U}\|_{H^m} < \varepsilon_0\}$ de \mathbf{U} et à valeurs dans un voisinage de 0 telle que

$$\left(\int_{\mathbb{R}} \mathbf{V}_t \cdot \mathbf{U}' = 0, \mathbf{V} \in \mathcal{V}_0, |t| < \alpha \right) \iff t = \tau(\mathbf{V}).$$

Ensuite, si $\min_{s \in \mathbb{R}} \|\mathbf{V} - \mathbf{U}_s\|_{H^m} = \|\mathbf{V} - \mathbf{U}_{s_0}\|_{H^m} < \varepsilon_0$, on a évidemment $\|\mathbf{V}_{-s_0} - \mathbf{U}\|_{H^m} < \varepsilon_0$ et l'on peut définir

$$\tau(\mathbf{V}) := \tau(\mathbf{V}_{-s_0}) - s_0,$$

de sorte que

$$\int \mathbf{V}_{\tau(\mathbf{V})} \cdot \mathbf{U}' = 0$$

puisque par définition de τ ,

$$\int \mathbf{V}_{-s_0 + \tau(\mathbf{V}_{-s_0})} \cdot \mathbf{U}' = 0.$$

Pour vérifier que $\tau(\mathbf{V}_r) = \tau(\mathbf{V}) - r$ quels que soient $r \in \mathbb{R}$ et $\mathbf{V} \in \mathcal{V}$, on observe que, si $r \in \mathbb{R}$ et $\mathbf{V} \in \mathcal{V}$, alors $\mathbf{V}_r \in \mathcal{V}$, et si $\min_{t \in \mathbb{R}} \|\mathbf{V}_r - \mathbf{U}_t\|$ est atteint en t_0 , $\min_{s \in \mathbb{R}} \|\mathbf{V} - \mathbf{U}_s\|$ est atteint en $s_0 = t_0 - r$. Par suite, si $\|\mathbf{V}_r - \mathbf{U}_{t_0}\|_{H^m} < \varepsilon_0$, on a d'une part $\tau(\mathbf{V}_r) = \tau(\mathbf{V}_{r-t_0}) - t_0$ et d'autre part $\tau(\mathbf{V}) = \tau(\mathbf{V}_{r-t_0}) - (t_0 - r)$. En faisant la différence on en déduit, $\tau(\mathbf{V}_r) = \tau(\mathbf{V}) - r$. ■

Théorème IV.1

On suppose que \mathbf{U} est une solution stationnaire de (IV.7) dans $H^m(\mathbb{R}; \mathbb{R}^n)$ avec $m \geq 1$, telle qu'il existe $\alpha > 0$ tel que

$$(\mathcal{H} - \sigma \mathcal{J})[\mathbf{u}] - (\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}] \geq \alpha \|\mathbf{u} - \mathbf{U}\|_{H^m}^2$$

pour tout $\mathbf{u} \in (\mathbf{U}')^\perp$ (l'orthogonal étant pris dans $L^2(\mathbb{R}; \mathbb{R}^n)$) tel que

$$\|\mathbf{u} - \mathbf{U}\|_{H^m} \leq \eta_0.$$

On suppose en outre que \mathcal{H} est continue dans $H^m(\mathbb{R}; \mathbb{R}^n)$ et que le problème de Cauchy pour (IV.7) est localement bien posé dans $H^m(\mathbb{R}; \mathbb{R}^n)$. Alors \mathbf{U} est une solution orbitalement stable de (IV.7).

Démonstration : C'est une adaptation du théorème de Lyapunov classique, utilisant le lemme IV.1 pour éliminer la « mauvaise » direction \mathbf{U}' liée à l'invariance par translation. On suppose sans perte de généralité $\eta_0 < \varepsilon_0$ (le ε_0 du lemme), et qu'il existe $T > 0$ tel que pour toute donnée initiale dans $H^m(\mathbb{R}; \mathbb{R}^n)$ à distance au plus η_0 de \mathbf{U} l'équation (IV.7) admet une solution $\mathbf{u} \in \mathcal{C}(0, T; H^m)$. Soit

$$a := (\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}] + \alpha \eta_0^2.$$

On a par hypothèse

$$\inf \left\{ (\mathcal{H} - \sigma \mathcal{J})[\mathbf{V}] ; \mathbf{V} \in (\mathbf{U}')^\perp, \|\mathbf{V} - \mathbf{U}\|_{H^m} = \eta_0 \right\} \geq a > (\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}].$$

Par continuité de $\mathcal{H} - \sigma \mathcal{J}$ et de $\mathbf{V} \mapsto \mathbf{V}_{\tau(\mathbf{V})}$, il existe $\eta \in]0, \eta_0[$ tel que pour tout $\mathbf{V} \in H^m$ tel que $\|\mathbf{V} - \mathbf{U}\|_{H^m} \leq \eta$ on ait

$$(\mathcal{H} - \sigma \mathcal{J})[\mathbf{V}] < a \quad \text{et} \quad \|\mathbf{V}_{\tau(\mathbf{V})} - \mathbf{U}\|_{H^m} < \eta_0.$$

Soit donc $\mathbf{u}_0 \in H^m$ tel que $\|\mathbf{u}_0 - \mathbf{U}\|_{H^m} \leq \eta$. Puisque $\eta < \eta_0$, l'équation (IV.7) admet une solution $\mathbf{u} \in \mathcal{C}(0, T; H^m)$ vérifiant $\mathbf{u}(0) = \mathbf{u}_0$, et l'on a pour tout $t \in [0, T]$,

$$\|\tilde{\mathbf{u}}(t) - \mathbf{U}\|_{H^m} < \eta_0,$$

où l'on a noté pour simplifier

$$\tilde{\mathbf{u}}(t) = \mathbf{u}(t)_{\tau(\mathbf{u}(t))}.$$

En effet, comme c'est vrai à $t = 0$, s'il existait $t_1 \in]0, T]$ tel que

$$\|\tilde{\mathbf{u}}(t_1) - \mathbf{U}\|_{H^m} \geq \eta_0,$$

par le théorème des valeurs intermédiaires et le principe de la borne supérieure il existerait $t_* \in]0, t_1]$ tel que

$$\|\tilde{\mathbf{u}}(t_*) - \mathbf{U}\|_{H^m} = \eta_0,$$

et pour tout $t \in [0, t_*[$,

$$\|\tilde{\mathbf{u}}(t) - \mathbf{U}\|_{H^m} < \eta_0.$$

Donc par définition de a on aurait

$$(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}(t_*)] = (\mathcal{H} - \sigma \mathcal{I})[\tilde{\mathbf{u}}(t_*)] \geq a,$$

tandis que pour tout $t \in [0, t_*[$,

$$(\mathcal{H} - \sigma \mathcal{I})[\tilde{\mathbf{u}}(t)] = (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}(t)] = (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}_0] < a.$$

Ceci contredirait la continuité de $t \mapsto (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}(t)]$.

On a donc pour tout $t \in [0, T]$,

$$(\mathcal{H} - \sigma \mathcal{I})[\tilde{\mathbf{u}}(t)] = (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}(t)] = (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}_0] < a \quad \text{et} \quad \|\tilde{\mathbf{u}}(t) - \mathbf{U}\|_{H^m} < \eta_0.$$

On peut par conséquent repartir de $\tilde{\mathbf{u}}(T)$ comme donnée initiale en procédant comme ci-dessus : ceci permet, en translatant la solution obtenue de $\tau(\mathbf{u}(T))$ en espace, de prolonger la solution du problème de départ à l'intervalle $[T, 2T]$ de telle façon que pour tout $t \in [T, 2T]$,

$$(\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}(t)] = (\mathcal{H} - \sigma \mathcal{I})[\mathbf{u}_0] < a \quad \text{et} \quad \|\mathbf{u}(t)_{\tau_1(t)} - \mathbf{U}\|_{H^m} < \eta_0,$$

où $\tau_1(t) = \tau(\mathbf{u}(T)) + \tau(\mathbf{u}(t)_{\tau(\mathbf{u}(T))})$. On montre ainsi successivement que u est définie pour tout $t \in \mathbb{R}$ et vérifie

$$\inf_{s \in \mathbb{R}} \|\mathbf{u}(t) - \mathbf{U}_s\|_{H^m} < \eta_0. \quad \blacksquare$$

Malheureusement, le théorème IV.1 ne s'applique pas si $\text{Hess}(\mathcal{H} - \sigma \mathcal{I})[\mathbf{U}]$ a une direction propre associée à une valeur propre négative, ce qui est généralement le cas pour les solitons !

Exemple. Reprenons l'équation de KdV (IV.2), pour laquelle

$$\mathbf{J} = 1, \quad H(u, u_x) = \frac{1}{2}u_x^2 + F(u), \quad I(u) = -\frac{1}{2}u^2,$$

où F est une primitive de f . La hessienne

$$\mathcal{L}^\sigma := \text{Hess}(\mathcal{H} - \sigma \mathcal{I})[U] = -\partial_{xx}^2 + f'(U) + \sigma$$

est un opérateur de *Sturm-Liouville*, ayant la dérivée du profil U dans son noyau. Or, si U est un soliton, U' s'annule une fois, et la théorie de Sturm-Liouville montre que 0 (valeur propre associée à U') n'est pas la plus petite valeur propre de \mathcal{L}^σ : il en existe exactement une strictement négative.

Cependant, la « mauvaise » direction peut dans certains cas être éliminée, en utilisant le fait que les solitons sont paramétrés par leur vitesse. C'est l'objet du résultat suivant.

Théorème IV.2 (*Grillakis-Shatah-Strauss*)

On suppose que le problème de Cauchy pour (IV.7) est localement bien posé dans $H^m(\mathbb{R}; \mathbb{R}^n)$ et que \mathcal{H} est continue dans $H^m(\mathbb{R}; \mathbb{R}^n)$. On suppose qu'il existe une famille de solitons \mathbf{U}^σ de vitesse $\sigma \in \Sigma$ (un intervalle ouvert), tels que, pour tout $\sigma \in \Sigma$,

$\mathcal{L}^\sigma := \text{Hess}(\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}^\sigma]$ a exactement une valeur propre négative, son noyau étant engendré par $\partial_x \mathbf{U}^\sigma$ et le reste de son spectre étant dans $]\alpha, +\infty[$ avec $\alpha > 0$. Si de plus

$$(IV.8) \quad \frac{d^2}{d\sigma^2}(\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}^\sigma] > 0$$

alors \mathbf{U}^σ est une solution orbitalement stable de (IV.7).

Ceci est en fait un cas particulier du théorème démontré par Grillakis, Shatah et Strauss, qui montrent aussi que la condition (IV.8) est nécessaire pour la stabilité des solitons *lorsque \mathcal{J} est surjectif* : ceci n'est clairement pas le cas lorsque $\mathcal{J} = D_x$. Cependant, cette difficulté a été levée dans le cas de (KdV) par Bona-Souganidis-Strauss.

Remarque IV.1

D'après l'équation de profil (IV.6),

$$\frac{d}{d\sigma}(\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}^\sigma] = -\mathcal{J}[\mathbf{U}^\sigma],$$

et donc

$$\frac{d^2}{d\sigma^2}(\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}^\sigma] = -\frac{d}{d\sigma}\mathcal{J}[\mathbf{U}^\sigma] = -\int \delta \mathcal{J}[\mathbf{U}^\sigma] \cdot \partial_\sigma \mathbf{U}^\sigma.$$

Or en dérivant (IV.6) par rapport à σ on voit que

$$\delta \mathcal{J}[\mathbf{U}^\sigma] = \mathcal{L}^\sigma \partial_\sigma \mathbf{U}^\sigma.$$

Par suite,

$$\frac{d^2}{d\sigma^2}(\mathcal{H} - \sigma \mathcal{J})[\mathbf{U}^\sigma] = -\langle \partial_\sigma \mathbf{U}^\sigma, \mathcal{L}^\sigma \partial_\sigma \mathbf{U}^\sigma \rangle.$$

La preuve du théorème est fondée sur l'idée que

$$\langle \mathbf{V}, \mathcal{L}^\sigma \mathbf{V} \rangle \geq 0 \quad \forall \mathbf{V} \in \delta \mathcal{J}[\mathbf{U}^\sigma]^\perp,$$

c'est-à-dire que \mathbf{U}^σ est un minimum local de $\mathcal{H} - \sigma \mathcal{J}$ sur la « variété » $\{\mathcal{J}[\mathbf{V}] = \mathcal{J}[\mathbf{U}^\sigma]\}$.

3 Compléments

- Multi-solitons (méthode de Hirota)
- Scattering/scattering inverse

Chapitre A

Rappels d'analyse fonctionnelle

L'un des piliers de l'analyse fonctionnelle est le théorème de Hahn-Banach, qui peut être énoncé sous diverses formes (voir [3, ch.1]). Nous l'utilisons dans ce cours essentiellement sous la forme suivante.

Théorème A.1 (*Hahn-Banach*)

Soit E un espace vectoriel normé, et G un sous-espace vectoriel de E . Si f est une application linéaire continue de G dans \mathbb{R} , elle se prolonge en une application linéaire continue de E dans \mathbb{R} , c'est-à-dire une forme linéaire continue, de norme

$$\sup\{|f(x)|; x \in G, \|x\|_E \leq 1\}.$$

Définition A.1

Soit A un opérateur linéaire de domaine \mathcal{D} , sous-espace d'un espace de Banach \mathcal{X} , à valeurs dans un espace de Banach \mathcal{Y} . On note \mathcal{X}^* le dual topologique de \mathcal{X} (c'est-à-dire l'espace des formes linéaires continues sur \mathcal{X} , muni de la norme subordonnée à celle de \mathcal{X}), de même que \mathcal{Y}^* est le dual topologique de \mathcal{Y} . On appelle adjoint de A l'opérateur non borné de domaine

$$\mathcal{D}(A^*) := \{\varphi \in \mathcal{Y}^*; \sup_{f \in \mathcal{D}, \|f\|_{\mathcal{X}} \leq 1} \langle \varphi, Af \rangle < +\infty\}$$

tel que

$$\langle A^*v, f \rangle = \langle v, Af \rangle, \quad \forall v \in \mathcal{D}(A^*), \forall f \in \mathcal{D}.$$

D'après le théorème de Hahn-Banach, ceci permet de définir A^*v comme une forme linéaire continue sur \mathcal{X} , c'est-à-dire un élément de \mathcal{X}^* , et ainsi A^* comme un opérateur non borné de $\mathcal{D}(A^*) \subset \mathcal{Y}^*$ dans \mathcal{X}^* . Si \mathcal{X} est en fait un espace de Hilbert, il s'identifie à \mathcal{X}^* par la formule de représentation de Riesz-Fréchet. On dit alors que A est autoadjoint si $A = A^*$, ce qui demande en fait deux choses :

1. égalité des domaines : $\mathcal{D}(A^*) = \mathcal{D}(A)$
2. symétrie : $\langle Av, f \rangle = \langle v, Af \rangle, \quad \forall v, f \in \mathcal{D}(A).$

De façon analogue, A est dit anti-autoadjoint si $A = -A^*$.

Chapitre B

Résumé d'analyse de Fourier

La transformation de Fourier est un opérateur linéaire continu

$$\begin{array}{ccc} \mathcal{F} : L^2(\mathbb{R}^d) & \rightarrow & L^2(\mathbb{R}^d) \\ u & \mapsto & \widehat{u}, \end{array}$$

où \widehat{u} est défini explicitement si $u \in L^1(\mathbb{R}^d)$ par la formule

$$\widehat{u}(\boldsymbol{\xi}) = \int_{\mathbb{R}^d} u(\mathbf{x}) e^{-i\boldsymbol{\xi} \cdot \mathbf{x}} d\mathbf{x}.$$

De plus \mathcal{F} est un isomorphisme, et sa réciproque est donnée par la *formule d'inversion de Fourier*, pour $U \in L^1(\mathbb{R}^d)$,

$$\mathcal{F}^{-1}(U)(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} U(\boldsymbol{\xi}) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\boldsymbol{\xi}.$$

Au facteur $(2\pi)^{d/2}$ près, \mathcal{F} est une isométrie. On a effet la *formule de Plancherel* : pour tout $u \in L^2(\mathbb{R}^d)$,

$$\|u\|_{L^2} = (2\pi)^{-d/2} \|\widehat{u}\|_{L^2}.$$

La classe de Schwartz des fonctions à décroissance rapide

$$\mathcal{S}(\mathbb{R}^d) = \left\{ f \in \mathcal{C}^\infty(\mathbb{R}^d); \forall \alpha \in \mathbb{N}^d, \forall \beta \in \mathbb{N}, \sup_{\mathbf{x} \in \mathbb{R}^d} (1 + \|\mathbf{x}\|)^\beta |\partial^\alpha f(\mathbf{x})| < +\infty \right\}$$

est invariante par transformation de Fourier. Ceci permet d'étendre cette transformation aux distributions tempérées

$$\begin{array}{ccc} \mathcal{F} : \mathcal{S}'(\mathbb{R}^d) & \rightarrow & \mathcal{S}'(\mathbb{R}^d) \\ T & \mapsto & \widehat{T}; \langle \widehat{T}, f \rangle = \langle T, \widehat{f} \rangle \text{ pour tout } f \in \mathcal{S}(\mathbb{R}^d). \end{array}$$

Remarque B.1

Si $f \in \mathcal{D}(\mathbb{R}^d)$ (espace des fonctions de classe \mathcal{C}^∞ à support compact), sa transformée de Fourier \widehat{f} se prolonge en une fonction analytique sur \mathbb{C}^d .

En effet, si $K = \text{supp}(f)$,

$$\widehat{f}(\boldsymbol{\xi}) = \int_K f(\mathbf{x}) e^{-i\boldsymbol{\xi} \cdot \mathbf{x}} d\mathbf{x}$$

est défini quel que soit $\boldsymbol{\xi} \in \mathbb{C}^d$ et hérite de l'analyticité de la fonction exponentielle. De plus, en notant $\|\boldsymbol{\xi}\| = \sqrt{\sum_{j=1}^d |\xi_j|^2}$, $\boldsymbol{\eta} = \text{Im}\boldsymbol{\xi}$ et

$$I_K(\boldsymbol{\eta}) = \max_{\mathbf{x} \in K} (\mathbf{x} \cdot \boldsymbol{\eta}),$$

on montre par intégrations par parties successives, l'inégalité

$$|\widehat{f}(\boldsymbol{\xi})| \leq \frac{1}{\|\boldsymbol{\xi}\|^{|\alpha|}} \|\partial^\alpha f\|_{L^1(K)} e^{I_K(\boldsymbol{\eta})},$$

pour tout d -uplet α . Par suite, pour tout $p \in \mathbb{N}^*$, il existe $C_p > 0$ tel que

$$|\widehat{f}(\boldsymbol{\xi})| \leq \frac{C_p}{(1 + \|\boldsymbol{\xi}\|)^p} e^{I_K(\boldsymbol{\eta})}$$

quel que soit $\boldsymbol{\xi} \in \mathbb{C}^d$. (Noter qu'en particulier si K est la boule de centre 0 et de rayon R , $I_K(\boldsymbol{\eta}) = R \|\boldsymbol{\eta}\|$.) Cette propriété caractérise en fait la transformée de Fourier des fonctions de classe \mathcal{C}^∞ à support inclus dans K : ceci est l'objet du *théorème de Paley-Wiener* énoncé ci-dessous.

Théorème B.1 (Paley–Wiener)

Si F est une fonction analytique sur \mathbb{C}^d pour laquelle il existe un compact K de \mathbb{R}^d tel que pour tout $p \in \mathbb{N}$, il existe $C_p > 0$ avec

$$|V(\boldsymbol{\xi})| \leq \frac{C_p}{(1 + \|\boldsymbol{\xi}\|)^p} \exp\left(\max_{\mathbf{x} \in K} (\mathbf{x} \cdot \text{Im}\boldsymbol{\xi})\right)$$

for all $\boldsymbol{\xi} \in \mathbb{C}^d$, alors $F|_{\mathbb{R}^d}$ est la transformée de Fourier d'une fonction $f \in \mathcal{D}(\mathbb{R}^d)$ à support dans K .

Théorème B.2 (Paley–Wiener–Schwartz)

Pour $u \in \mathcal{E}'(\mathbb{R}^d)$ (ensemble des distributions à support compact), la transformée de Fourier s'étend en une fonction analytique sur \mathbb{C}^d par la formule

$$U(\boldsymbol{\xi}) = \langle u, e^{-i\boldsymbol{\xi} \cdot \cdot} \rangle.$$

De plus si $K = \text{Supp} u$, il existe un entier p et $C > 0$ tels que

$$|U(\boldsymbol{\xi})| \leq C (1 + \|\boldsymbol{\xi}\|)^p \exp\left(\max_{\mathbf{x} \in K} (\mathbf{x} \cdot \text{Im}\boldsymbol{\xi})\right)$$

for all $\boldsymbol{\xi} \in \mathbb{C}^d$. Réciproquement, si U est une fonction analytique sur \mathbb{C}^d pour laquelle il existe un compact K de \mathbb{R}^d , un entier p , et une constante C tels que l'estimation ci-dessus soit satisfaite, alors U est la transformée de Fourier d'une distribution à support dans K .

En résumé, la transformation de Fourier échange les propriétés de localisation et de régularité.

Principe d'incertitude de Heisenberg

Théorème B.3

Soit $f \in H^1(\mathbb{R})$ telle que $x \mapsto xf(x)$ soit de carré intégrable. On définit la dispersion de f comme le nombre

$$D_f := \inf_{x \in \mathbb{R}} \left(\int (y - x)^2 |f(y)|^2 dy \right)^{1/2}.$$

Alors le produit $D_f D_{\hat{f}}$ est minoré par une constante strictement positive (indépendante de f !).

Démonstration : Ce résultat est une conséquence la formule de Plancherel et de l'inégalité de Cauchy-Schwarz. En effet, si une fonction f vérifie les hypothèses du théorème, l'infimum dans la définition de D_f et dans celle de $D_{\hat{f}}$ est atteint, car les fonctions $x \mapsto \int (y - x)^2 |f(y)|^2 dy$ et $\xi \mapsto \int (\zeta - \xi)^2 |\hat{f}(\zeta)|^2 d\zeta$ sont convexes et minorées par zéro. Sans perte de généralité on peut supposer qu'ils sont atteints en zéro, quitte à changer f en $y \mapsto e^{-i\xi_0 y} f(y + x_0)$ (dont la dispersion est atteinte en $x = 0$ si celle de f l'est en x_0 , de même que la dispersion de sa transformée de Fourier $\zeta \mapsto e^{i(\zeta + \xi_0)} \hat{f}(\zeta + \xi_0)$ est atteinte en $\zeta = 0$ si celle de \hat{f} est atteinte en ξ_0). Ainsi, on a

$$D_f^2 = \int y^2 |f(y)|^2 dy, \quad D_{\hat{f}}^2 = \int \zeta^2 |\hat{f}(\zeta)|^2 d\zeta = 2\pi \int |f'(y)|^2 dy$$

d'après la formule de Plancherel et l'identité $\mathcal{F}(f')(\zeta) = \zeta \hat{f}(\zeta)$. Or, d'après la inégalité de Cauchy-Schwarz,

$$\left(\int y^2 |f(y)|^2 dy \right)^{1/2} \left(\int |f'(y)|^2 dy \right)^{1/2} \geq \left| \int y f(y) f'(y) dy \right| = \frac{1}{2} \int f(y)^2 dy$$

par intégration par parties. Ceci prouve l'inégalité

$$D_f D_{\hat{f}} \geq \sqrt{\frac{\pi}{2}}.$$

(La valeur effective du minorant dépend de la convention choisie pour la définition de la transformation de Fourier.) ■

Transformées de Fourier classiques

- **Gaussiennes :** Pour $\alpha > 0$,

$$\mathcal{F}^{-1}(\xi \mapsto e^{-\alpha \|\xi\|^2})(\mathbf{x}) = \left(\frac{1}{4\alpha\pi} \right)^{d/2} e^{-\frac{\|\mathbf{x}\|^2}{4\alpha}}.$$

- **Noyau de Schrödinger :** Pour $h > 0$,

$$\mathcal{F}^{-1}(\xi \mapsto e^{ih\|\xi\|^2})(\mathbf{x}) = \left(\frac{1}{4h\pi} \right)^{d/2} e^{-id\pi/4} e^{i\frac{\|\mathbf{x}\|^2}{4h}}.$$

Cette formule se démontre en utilisant la précédente et le principe du prolongement analytique pour les fonctions holomorphes. En effet, quelle que soit la fonction test $\phi \in \mathcal{S}(\mathbb{R}^d)$, les fonctions de la variable complexe

$$z \mapsto \left\langle e^{-z\|\cdot\|^2}, \hat{\phi} \right\rangle$$

$$z \mapsto \left(\frac{1}{4z\pi} \right)^{d/2} \left\langle e^{-\frac{\|\cdot\|^2}{4z}}, \phi \right\rangle$$

coïncident sur l'axe des nombres réels positifs et sont holomorphes dans le demi-plan $\{\operatorname{Re} z > 0\}$ (on prend la détermination de partie réelle positive de la racine carrée), donc elles coïncident dans ce demi-plan. En passant à la limite $\operatorname{Re} z \searrow 0$ on obtient la formule voulue.

- **Noyau de Green des ondes en dimension 3 :** Pour $r > 0$,

$$\mathcal{F}^{-1} \left(\xi \mapsto \frac{\cos(r\|\xi\|)}{\|\xi\|^2} \right) (\mathbf{x}) = \frac{H(\|\mathbf{x}\| - r)}{4\pi\|\mathbf{x}\|},$$

où H est la fonction de Heaviside. Pour montrer cette formule on peut commencer par calculer la transformée de Fourier de la fonction $\mathbf{x} \mapsto \frac{\chi_R(\|\mathbf{x}\|)}{\|\mathbf{x}\|}$, où χ_R est la fonction caractéristique de $]r, R[$. Elle est nécessairement radiale, et en utilisant les coordonnées sphériques on a

$$\begin{aligned} \mathcal{F} \left(\mathbf{x} \mapsto \frac{\chi_R(\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) (0, 0, \xi_3) &= \int_r^R \int_0^{2\pi} \int_0^\pi e^{-i\xi_3 \rho \cos \theta} \rho \sin \theta \, d\rho \, d\varphi \, d\theta \\ &= \frac{4\pi}{\xi_3^2} (\cos(r\xi_3) - \cos(R\xi_3)), \end{aligned}$$

d'où

$$\mathcal{F} \left(\mathbf{x} \mapsto \frac{\chi_R(\|\mathbf{x}\|)}{\|\mathbf{x}\|} \right) (\xi) = \frac{4\pi}{\|\xi\|^2} (\cos(r\|\xi\|) - \cos(R\|\xi\|)).$$

Il reste ensuite à montrer que cette fonction tend *au sens des distributions* vers la fonction $\xi \mapsto \frac{4\pi \cos(r\|\xi\|)}{\|\xi\|^2}$ lorsque $R \rightarrow +\infty$. Soit $\phi \in \mathcal{S}(\mathbb{R}^3)$. On a

$$\begin{aligned} \int \frac{\cos(R\|\xi\|)}{\|\xi\|^2} \phi(\xi) \, d\xi &= \\ \int_0^{+\infty} \int_0^{2\pi} \int_0^\pi \cos(\rho R) \phi(\rho \sin \theta \cos \varphi, \rho \sin \theta \sin \varphi, \rho \cos \theta) \sin \theta \, d\rho \, d\varphi \, d\theta. \end{aligned}$$

En intégrant par parties en ρ , on voit que

$$\left| \int \frac{\cos(R\|\xi\|)}{\|\xi\|^2} \phi(\xi) \, d\xi \right| \lesssim \frac{1}{R} \int \frac{\|\nabla \phi(\xi)\|}{\|\xi\|^2} \, d\xi \rightarrow 0$$

lorsque $R \rightarrow +\infty$.

Espaces de Sobolev construits sur $L^2(\mathbb{R}^d)$ Pour tout $s > 0$, on définit

$$\lambda^s(\xi) = (1 + \|\xi\|^2)^{s/2},$$

et l'espace

$$H^s(\mathbb{R}^d) = \{ f \in L^2(\mathbb{R}^d); \lambda^s \widehat{f} \in L^2(\mathbb{R}^d) \}$$

muni de la norme

$$\|f\|_{H^s(\mathbb{R}^d)} = \|\lambda^s \widehat{f}\|_{L^2(\mathbb{R}^d)}.$$

C'est un espace de Hilbert. D'après la formule

$$\widehat{\partial^\alpha f(\boldsymbol{\xi})} = i^{|\alpha|} \boldsymbol{\xi}^\alpha \widehat{f(\boldsymbol{\xi})},$$

on voit que pour tout $p \in \mathbb{N}^*$, $H^p(\mathbb{R}^d)$ contient les fonctions p fois dérivables dont toutes les dérivées jusqu'à l'ordre p sont de carré intégrable. En fait, $H^p(\mathbb{R}^d)$ est exactement l'ensemble des *distributions tempérées* dont toutes les dérivées jusqu'à l'ordre p sont des fonctions appartenant à $L^2(\mathbb{R}^d)$. Et la norme sur $H^p(\mathbb{R}^d)$ est équivalente à la norme $\|\cdot\|$ définie par

$$\|u\|_{H^p(\mathbb{R}^d)}^2 = \|u\|_{L^2(\mathbb{R}^d)}^2 + \sum_{1 \leq |\alpha| \leq p} \|\partial^\alpha u\|_{L^2(\mathbb{R}^d)}^2.$$

On définit également les espaces de Sobolev homogènes

$$\dot{H}^s(\mathbb{R}^d) = \{f \in L^2(\mathbb{R}^d); \boldsymbol{\xi} \mapsto \|\boldsymbol{\xi}\|^s \widehat{f}(\boldsymbol{\xi}) \in L^2(\mathbb{R}^d)\}$$

munis de la norme

$$\|f\|_{\dot{H}^s(\mathbb{R}^d)} = \|\boldsymbol{\xi} \mapsto \|\boldsymbol{\xi}\|^s \widehat{f}(\boldsymbol{\xi})\|_{L^2(\mathbb{R}^d)}.$$

On a clairement $H^s(\mathbb{R}^d) \subset \dot{H}^s(\mathbb{R}^d)$ avec

$$\|f\|_{\dot{H}^s(\mathbb{R}^d)} \lesssim \|f\|_{H^s(\mathbb{R}^d)},$$

et

$$\|f\|_{\dot{H}^s(\mathbb{R}^d)} \lesssim \|\nabla f\|_{H^{s-1}(\mathbb{R}^d)}.$$

De plus,

$$\|f\|_{H^s(\mathbb{R}^d)} \approx \|f\|_{L^2(\mathbb{R}^d)} + \|f\|_{\dot{H}^s(\mathbb{R}^d)}, \quad s > 0.$$

Injections de Sobolev :

$$H^s(\mathbb{R}^d) \hookrightarrow L^\infty(\mathbb{R}^d), \quad s > d/2.$$

$$\dot{H}^s(\mathbb{R}^d) \hookrightarrow L^q(\mathbb{R}^d), \quad s \in]0, d/2[, \quad \frac{1}{q} = \frac{1}{2} - \frac{s}{d}.$$

Bibliographie

- [1] Sylvie Benzoni-Gavage and Denis Serre. *Multidimensional hyperbolic partial differential equations*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, Oxford, 2007. First-order systems and applications.
- [2] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.
- [3] Haïm Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [4] Jean Dieudonné. *Calcul infinitésimal*. Hermann, Paris, 1968.
- [5] P. G. Drazin and R. S. Johnson. *Solitons : an introduction*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 1989.
- [6] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian methods in the theory of solitons*. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1987. Translated from the Russian by A. G. Reyman [A. G. Reĭman].
- [7] J. Ginibre and G. Velo. Smoothing properties and retarded estimates for some dispersive evolution equations. *Comm. Math. Phys.*, 144(1) :163–188, 1992.
- [8] Lars Hörmander. *The analysis of linear partial differential operators. I*, volume 256 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1990. Distribution theory and Fourier analysis.
- [9] T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. In *Spectral theory and differential equations (Proc. Sympos., Dundee, 1974 ; dedicated to Konrad Jörgens)*, pages 25–70. Lecture Notes in Math., Vol. 448. Springer, Berlin, 1975.
- [10] T. Kato. On the Cauchy problem for the (generalized) Korteweg-de Vries equation. In *Studies in applied mathematics*, volume 8 of *Adv. Math. Suppl. Stud.*, pages 93–128. Academic Press, New York, 1983.
- [11] Alan C. Newell. *Solitons in mathematics and physics*, volume 48 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1985.
- [12] Peter J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.
- [13] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

- [14] Terence Tao. *Nonlinear dispersive equations*, volume 106 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. Local and global analysis.
- [15] G. B. Whitham. *Linear and nonlinear waves*. Pure and Applied Mathematics. John Wiley & Sons Inc., New York, 1999. Reprint of the 1974 original, A Wiley-Interscience Publication.

Index

- équation
 - d’Airy, 3
 - de Boussinesq, 3
 - de Hamilton-Jacobi, 8
 - de Klein-Gordon, 3
 - de la chaleur, 7
 - de Schrödinger, 3
 - de sine-Gordon, 25
 - de transport, 3
 - des ondes, 3
 - eikonale, 8
- stable
 - solution asymptotiquement, 65
- a priori
 - estimation, 16
- Airy
 - équation d’, 3
 - fonction d’, 12
- anti-autoadjoint
 - opérateur, 71
- autoadjoint
 - opérateur, 71
- Boussinesq
 - équation de, 3
 - système de, 27
- caractéristiques
 - méthode des, 9
- Casimir
 - fonction, 53
- chaleur
 - équation de la, 7
 - noyau de la, 13
- champ de vecteurs, 44
 - généralisé, 48
 - généralisé évolutionnaire, 48
 - généralisé prolongé, 48
 - Hamiltonien, 52
 - Hamiltonien généralisé, 57
 - prolongé, 46
- crochet
 - de Poisson, 53, 55
 - crochet de Lie, 49
- d’Alembert
 - formule de, 14
- densité
 - conservée, 51
- diffusion, 7, 13
- dispersion, 13
 - relation de, 5
- dispersive
 - EDP, 7
- Duhamel
 - formule de, 18
 - opérateur de, 35
- estimation
 - a priori, 16
 - de dispersion, 17
- Euler
 - opérateur d’, 54
- flux, 51
- fonctionnelle, 54
- fondamentale
 - solution, 11
- formule
 - de Duhamel, 18
 - d’inversion de Fourier, 73
 - de d’Alembert, 14
 - de Kirchhoff, 14
 - de Plancherel, 73
- Fourier
 - formule d’inversion de, 73
 - transformation de, 73
- générateur infinitésimal

- d'un groupe de transformations à un paramètre, 44
- gradient variationnel, 54
- Green
 - fonction de, 11
 - noyau, 11
- groupe
 - de symétrie, 43
 - de transformations, 43
 - à un paramètre, 44
 - ligne de, 7
 - vitesse de, 7
- Hahn-Banach
 - théorème de, 71
- Hamiltonien, 52
- Heaviside
 - fonction de, 23
- homocline
 - orbite, 63
- Huygens
 - principe de, 14
- hyperbolique
 - EDP, 5
- intégrale première, 51
- interpolation, 16
 - bilinéaire, 38
- Jacobi
 - identité de Jacobi, 53
- Kirchhoff
 - formule de, 14
- lem :KPV, 39
- Lie
 - crochet de, 49
- Lyapunov
 - fonctionnelle de, 66
- m-quasi-accrétif
 - opérateur, 33
- nombre d'onde
 - local, 8
- noyau
 - de Green, 11
- onde
 - longueur d', 5
 - nombre d', 5
 - plane, 4
 - train d', 8
- OPPM, 5
- orbitale
 - stabilité, 66
- Paley-Wiener
 - théorème de , 74
- phase, 5
 - ligne de, 7
 - plane de, 63
 - portrait de, 63
 - vitesse de, 7
- Plancherel
 - formule de, 73
- Poisson
 - crochet de, 53, 55
- portrait de phase, 63
- profil, 61
- pulsation, 5
- quasi-linéaire
 - EDP, 27
- sécante hyperbolique, 62
- Schwartz
 - classe de, 73
- semi-groupe, 17
 - de contractions, 30
- semi-linéaire
 - EDP, 25
- Sobolev
 - espace de , 76
 - injections de, 77
- Sobolev homogène
 - espace de , 77
- solution
 - classique, 27
 - faible, 28
 - forte, 28
 - mild, 18
- sous-critique
 - espace, 26
- stable
 - solution, 65
 - solution orbitalement, 66

- Sturm-Liouville
 - opérateur de, 69
- sur-critique
 - espace, 26
- symbole, 4
 - principal, 4
- thm :KPV, 40
- unitaire
 - groupe, 18
- vitesse
 - de groupe, 7
 - de phase, 7

Table des matières

I	EDP dispersives linéaires	3
1	Exemples et définitions de base	3
2	Dispersion	5
2.1	Dimension un	5
2.2	Dimension quelconque	10
3	Formules de résolution	11
4	Estimations <i>a priori</i>	15
4.1	Conservation de l'énergie	15
4.2	Estimations de dispersion	16
4.3	Estimations de de Strichartz	17
II	EDP dispersives non-linéaires	25
1	Exemples et vocabulaire	25
1.1	NLS et NLW	25
1.2	KdV et Boussinesq	27
2	Problème de Cauchy	27
2.1	Les différentes notions de solutions	27
2.2	Méthode des semi-groupes	29
2.3	Solutions moins régulières	34
2.4	Solutions globales	40
III	Symétries et lois de conservation	43
1	Groupes de symétrie des EDP	43
2	Lois de conservation	51
3	Formalisme Hamiltonien et théorème de Noether	52
3.1	EDO	52
3.2	EDP	54
IV	Solitons	61
1	Existence des solitons	61
2	Stabilité des solitons	65
3	Compléments	70
A	Rappels d'analyse fonctionnelle	71
B	Résumé d'analyse de Fourier	73
	Bibliographie	78

